Universidade de Aveiro Departamento de Química
2015

Joana Isabel Monteiro Pinto

Gravidez saudável e doenças pré-natais estudadas por metabolómica de plasma sanguíneo

Healthy pregnancy and prenatal disorders followed by blood plasma metabolomics

Universidade de Aveiro Departamento de Química
2015

Joana Isabel Monteiro Pinto

Gravidez saudável e doenças pré-natais estudadas por metabolómica de plasma sanguíneo

Healthy pregnancy and prenatal disorders followed by blood plasma metabolomics

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Bioquímica, realizada sob a orientação científica da Doutora Ana Maria Pissarra Coelho Gil, Professora Associada com Agregação do Departamento de Química da Universidade de Aveiro e da Doutora Maria do Rosário Gonçalves dos Reis Marques Domingues, Professora Auxiliar com Agregação do Departamento de Química da Universidade de Aveiro

Apoio financeiro da Fundação para a Ciência e Tecnologia (FCT) - bolsa de Investigação FCT SFRH/BD/73343/2010 financiada pelo Programa Operacional Potencial Humano (POPH) e projecto PTDC/QUI/66523/2006 financiado pelo Fundo Europeu de Desenvolvimento Regional (FEDER) e pelo Programa Operacional Factores de Competitividade (COMPETE); da Universidade de Aveiro - Centro de Investigação em Materiais Cerâmicos e Compósitos (CICECO) (FCT UID/CTM/50011/2013), financiado por fundos nacionais da FCT/Ministério da Educação e Ciência e co-financiado pelo FEDER sob o Acordo de Parceria PT2020, e Química Orgânica, Produtos Naturais e Agroalimentares (QOPNA) financiado por fundos da FCT (Pest-C/QUI/UI0062/2013); da empresa Bruker BioSpin GmbH; da Rede Nacional de RMN (RNRMN), suportada com fundos da FCT; e da Rede Nacional de Espectrometria de Massa (RNEM).

Dedico este trabalho à minha mãe pelo incondicional
apoio e paciência em todos os momentos.

o júri

presidente

Professor Doutor Helmuth Malonek
Professor Catedrático do Departamento de Matemática da Universidade de Aveiro

Doutor John Griffith Jones
Investigador Principal do Centro de Neurociências e Biologia Celular da Universidade de Coimbra

Doutora Iola Duarte
Investigadora Principal do Centro de Investigação em Materiais Cerâmicos e Compósitos (CICECO) da Universidade de Aveiro

Doutora Maria Paula do Amaral Alegria Guedes de Pinho
Investigadora Auxiliar do Laboratório Associado UCIBIO@REQUIMTE, Laboratório de Toxicologia do Departamento de Ciências Biológicas da Faculdade de Farmácia da Universidade do Porto

Drª Maria do Céu Almeida
Assistente Hospitalar Sénior (Chefe de Serviço) na especialidade de Ginecologia/Obstetrícia e Directora do Serviço de Obstetrícia B da Maternidade Bissaya Barreto, Centro Hospitalar e Universitário de Coimbra

Doutora Ana Maria Pissarra Coelho Gil
Professora Associada com Agregação do Departamento de Química da Universidade de Aveiro

Acknowledgements

I would like to express my sincere gratitude to may supervisors Dr. Ana M. Gil and Dr. Rosário M. Domingues for giving me the opportunity to work with them in the field of metabolomics applied to prenatal research. I would like to thank them for all the guidance, support and the knowledge they passed to me.

I want to acknowledge the medical team involved in this research project, particularly Dr. Maria do Céu Almeida, Dr. Eulália Galhano and Nurse Cristina Pita (of the Maternidade Bissaya Barreto, Coimbra), as well as Dr. Isabel Marques Carreira (of the Cytogenetics and Genomics Laboratory, Faculty of Medicine, and CNC.IBILI, University of Coimbra, and CIMAGO Center for Research in Environment, Genetics and Oncobiology, Coimbra). I also acknowledge the Centro Hospitalar e Universitário de Coimbra (CHUC) for accommodating this research project.

I also express my gratitude to Dr. Iola Duarte and Dr. Brian Goodfellow (University of Aveiro) for their help with the NMR experiments, and Dr. António Barros (University of Aveiro) for their help and advice on multivariate analysis. A special thanks is due to Gonçalo Graça, from whom I received my initial training in samples preparation and NMR acquisition, and to Sílvia Diaz for sharing the NMR data of maternal urine used in this work and for her help and fellowship over these years. Lastly, I want to acknowledge Lara Almeida, for her contribution to this work through plasma lipid extraction, and to Elisabete Maciel and Tânia Melo for their contribution to MS acquisition.

I would like to acknowledge Prof. Stefano Caldarelli and his team (Institut des Sciences Moléculaires de Marseille, Aix Marseille Université, France), particularly Dr. Mehdi Yemloul, for providing access and training related to Multiple Quantum NMR experiments. I also want to express my gratitude to Dr. Laetitia Shintu, Dr. Rosy Rosas and Lamya Rezig for the assistance provided and for receiving me so well in the three months I spent in France. I also thank Dr. Manfred Spraul (Bruker Biopsin, Germany) for providing access to spectral databases.

I acknowledge the Portuguese National NMR Network (RNRMN), supported with funds from the Foundation of Science and Technology (FCT), the Portuguese Mass Spectrometry Network (RNEM), and financial support from the European regional Development Fund through the Competitive Factors Thematic Operational Program, FCT research grant (SFRH/BD/73343/2010), CICECO, QOPNA, and University of Aveiro, for providing the means necessary to successfully conclude this work.

During this journey, I have had the opportunity to work with an extraordinary group to which I deeply thank for their friendship, motivation and laugh moments particularly to Cláudia Rocha, Gonçalo Graça, Inês Lamego, Joana Carrola, Joana Marques, João Rodrigues, Sílvia Diaz and Susana Aveiro. A special thanks is due to Elisabete Maciel who has been a friend during good and bad times.

I am especially grateful to Nuno for his companionship, support and patience over the years.

Finally, I deeply thank my family, especially my mother, for their unconditional love, patience and support.

palavras-chave

Plasma sanguíneo materno, extractos lipídicos, gravidez, doenças pré-natais, metabolómica, espectroscopia de ressonância magnética nuclear (RMN), cromatografia liquida hidrofílica acoplada a espectrometria de massa (HILIC-LC/MS), análise multivariada (MVA)

resumo

O trabalho apresentado nesta tese teve como principal objetivo investigar o impacto da gravidez saudável e algumas doenças pré-natais no metaboloma e lipidoma de plasma sanguíneo materno, com vista à definição de novos biomarcadores para a previsão e diagnóstico não invasivos daquelas doenças.

O Capítulo 1 descreve a perspectiva atual e os desafios das doenças pré-natais mais relevantes, assim como a estratégia metabolómica e estado da arte na investigação pré-natal. Todos os detalhes experimentais do trabalho realizado estão descritos no Capítulo 2, incluindo as condições de amostragem, recolha e preparação das amostras, bem como os protocolos de aquisição e análise dos dados.

No Capítulo 3 descreve-se o metaboloma e lipidoma de plasma detectados por RMN 1D e 2D. Neste capítulo, a utilização de espectroscopia de RMN de quantum-múltiplo foi explorada, pela primeira vez, para caracterização de misturas lipídicas complexas.

O Capítulo 4 contribui para colmatar algumas falhas no conhecimento sobre a degradibilidade do plasma humano durante o manuseamento da amostra e armazenamento, e a importância de condições de colheita como o jejum. A utilização de tubos de colheita com heparina não mostrou interferência do polissacarídeo nos espectros conservando-se toda a informação espectral, enquanto que os tubos com EDTA deram origem a sinais interferentes provenientes do EDTA livre e complexado com Ca^{2+}/Mg^{2+}, cujo impacto na análise metabolómica é discutido. Relativamente à estabilidade do plasma à temperatura ambiente, foram observadas alterações nas lipoproteínas e compostos de colina a partir de 2.5 horas, enquanto que o armazenamento a -20ºC mostrou ser adequado até 7 dias, sendo o armazenamento a -80ºC aconselhado, particularmente para períodos de tempo longos (pelo menos até 2.5 anos). Relativamente aos ciclos de congelação-descongelação, não se aconselham mais de 3 ciclos consecutivos, enquanto que o efeito da colheita das amostras em não-jejum (em vez de jejum) foi considerado aceitável.

O Capítulo 5 apresenta o primeiro estudo de metabolómica por RMN do plasma materno ao longo da gravidez, incluindo correlação entre plasma e urina. Algumas das alterações metabólicas observadas confirmaram efeitos metabólicos conhecidos, tendo outras sido observadas pela primeira vez sugerindo alterações no metabolismo energético, na microflora bacteriana (citrato, lactato e dimetil sulfona) e na taxa de filtração glomerular (creatina e creatinina). Os estudos de correlação revelaram aspetos metabólicos específicos das lipoproteínas/proteínas com impacto no metaboloma excretado.

No Capítulo 6 descreve-se o impacto das doenças cromossómicas (CD), incluindo Trissomia 21 (T21) no metaboloma e lipidoma de plasma materno.

Obtiveram-se elevadas taxas de classificação para CD (88-89%) e T21 (85-92%) no 1º e 2º trimestres baseadas na seleção de variáveis dos dados de RMN. A correlação de plasma e urina revelou novos desvios metabólicos, nomeadamente no metabolismo das lipoproteínas de baixa densidade e de muito baixa densidade (LDL+VLDL), dos açúcares e da microflora bacteriana. As alterações observadas no perfil de fosfolípidos do plasma, nomeadamente das fosfatidilcolinas, foram confirmadas e caracterizadas por cromatografia liquida hidrofílica acoplada a espetrometria de massa (HILIC-LC/MS).

No Capítulo 7 apresentam-se os resultados obtidos na prospecção de biomarcadores metabólicos de diabetes mellitus gestacional (GDM) pré- e pós-diagnóstico por metabolómica de RMN de plasma materno do 2º trimestre. Observaram-se alterações metabólicas com poder de previsão de GDM, nomeadamente um aumento no colesterol, ácidos gordos, triglicerídeos e pequenas variações metabólicas na glucose, aminoácidos, betaína, ureia, creatina e metabolitos relacionados com a microflora bacteriana. O grupo de GDM pós-diagnóstico foi bem classificado utilizando como biomarcador um conjunto de 26 ressonâncias do espectro de plasma correspondendo a lípidos e 10 metabolitos de baixo peso molecular, sugerindo-se a possibilidade de usar este marcador conjunto na gestão clínica da GDM.

O Capítulo 8 descreve os resultados obtidos para as doenças pré-natais que mostraram ter um menor impacto no metaboloma de plasma materno, nomeadamente as malformações fetais (FM), e os estados de pré-diagnóstico da rutura prematura das membranas (PROM), parto pré-termo (PTD) e pré-eclampsia.

Finalmente, no Capítulo 9 são descritas as conclusões gerais e perspetivas futuras no contexto desta tese, realçando-se como este trabalho contribui para o novo conhecimento dos mecanismos das doenças pré-natais e possíveis biomarcadores para a sua previsão e diagnóstico.

keywords

Maternal blood plasma, lipid extracts, pregnancy, prenatal disorders, metabolomics, nuclear magnetic resonance spectroscopy (NMR), hydrophilic interaction liquid chromatography-mass spectrometry (HILIC-LC/MS), multivariate analysis (MVA)

abstract

The work presented in this thesis aimed to investigate the impact of healthy pregnancy and selected prenatal disorders on the metabolome and lipidome of maternal blood plasma, in order to define new potential biomarkers for non-invasive prediction and diagnosis.

Chapter 1 describes the present status and challenges of the clinically relevant prenatal disorders, along with a presentation of the metabolomics strategy applied and the state of the art of metabolomics in prenatal research. All experimental details are described in Chapter 2, comprising sample metadata, sample collection and preparation, data acquisition protocols and data analysis procedures.

The plasma metabolome and lipidome viewed by 1D and 2D NMR experiments are presented in Chapter 3. In this chapter, the use of Multiple Quantum NMR spectroscopy was explored, for the first time, for assignment of complex lipid mixtures. Chapter 4 contributes to filling in some existing gaps regarding human plasma degradability during handling and storage, as well as the importance of fasting conditions at collection. The use of heparin collection tubes resulted in no interference of the polysaccharide and full conservation of spectral information, while EDTA tubes produced a number of interfering signals from free and Ca^{2+}/Mg^{2+} complexed EDTA, the impact of which on metabolomic analysis is discussed. Regarding temperature stability, large changes in lipoproteins and choline compounds were observed in plasma kept at room temperature for $\geq$ 2.5 hours, whereas short-term storage at -20ºC was found suitable up to 7 days, with storage at -80ºC being recommended, particularly for long-term periods (at least up to 2.5 years). Regarding freeze-thaw cycles, no more than 3 consecutive cycles were found advisable, while the use of non-fasting conditions (instead of fasting) was found acceptable.

Chapter 5 presents the first NMR metabolomics study of maternal plasma throughout pregnancy, including correlation between plasma and urine metabolites. Some of the metabolic alterations observed confirmed known metabolic effects, while novel changes were observed, suggesting adjustments in energy and gut microflora metabolisms (citrate, lactate and dimethyl sulfone) and alterations in glomerular filtration rate (creatine and creatinine). Correlations studies unveiled specific lipoprotein/protein metabolic aspects of healthy pregnancy with impact on the excreted metabolome, providing further understanding of pregnancy metabolism.

In Chapter 6, the impact of prenatal disorders on maternal plasma metabolome and lipidome is described for fetal chromosomal disorders (CD), including Trisomy 21 (T21). High classification rates were obtained for CD (88-89%) and T21 (85-92%) in 1^{st} and 2^{nd} trimesters, based on variable selection of NMR data. In addition, novel metabolic deviations were found through plasma/urine correlations, namely in low density and very low density lipoproteins

(LDL+VLDL), sugar and gut microflora metabolisms. Changes in plasma phospholipid profile, namely in phosphatidylcholines, were further confirmed and characterised by hydrophilic interaction liquid chromatography-mass spectrometry (HILIC-LC/MS).

In Chapter 7, metabolic biomarkers of pre- and post-diagnosis GDM were sought by NMR metabolomics of whole maternal plasma and plasma lipid profile in the 2nd trimester. Metabolic alterations found to be predictive of GDM comprised increases in cholesterol, fatty acids, triglycerides and small metabolites changes in glucose, amino acids, betaine, urea, creatine and metabolites related with gut microflora. Post-diagnosis GDM was successfully classified using a 26-resonance plasma biomarker corresponding to 10 metabolites and lipids, advancing the possibility of using a multi-metabolite biomarker as a complementary tool in the clinical management of GDM.

Chapter 8 describes the results obtained for prenatal disorders shown to have lower impact on maternal plasma metabolome, namely diagnosed fetal malformations and pre-diagnosis premature rupture of membranes, preterm delivery and preeclampsia.

Finally, Chapter 9 describes the general conclusions and future perspectives in the context of this thesis, highlighting how this work contributes with new knowledge on prenatal disease mechanisms and possible biomarkers for prenatal diagnosis and prediction methods.

List of publications including the work presented in this thesis

Papers in peer-reviewed journals:

Diaz, S. O.; Pinto, J.; Graça, G.; Duarte, I. F.; Barros, A. S.; Galhano, E.; Pita, C.; Almeida, M. do C.; Goodfellow, B. J.; Carreira, I. M.; Gil, A. M. Metabolic Biomarkers of Prenatal Disorders: An Exploratory NMR Metabonomics Study of Second Trimester Maternal Urine and Blood Plasma. *Journal of Proteome Research* **2011**, 10, 3732–3742.

Graça, G.; Diaz, S. O.; Pinto, J.; Barros, A. S.; Duarte, I. F.; Goodfellow, B. J.; Galhano, E.; Pita, C.; Almeida, M. do C.; Carreira, I. M.; Gil, A. M. Can Biofluids Metabolic Profiling Help to Improve Healthcare during Pregnancy? *Spectroscopy: An International Journal* **2012**, 27, 515–523.

Pinto, J.; Domingues, M. R. M.; Galhano, E.; Pita, C.; Almeida, M. do C.; Carreira, I. M.; Gil, A. M. Human Plasma Stability during Handling and Storage: Impact on NMR Metabolomics. *The Analyst* **2014**, 139, 1168–1177.

Pinto, J.; Maciel, E.; Melo, T. S.; Domingues, M. R. M.; Galhano, E.; Pita, C.; Almeida, M. do C.; Carreira, I. M.; Gil, A. M. Maternal Plasma Phospholipids Are Altered in Trisomy 21 Cases and prior to Preeclampsia and Preterm Outcomes. *Rapid Communications in Mass Spectrometry* **2014**, 28, 1635–1638.

Pinto, J.; Barros, A. S.; Domingues, M. R. M.; Goodfellow, B. J.; Galhano, E.; Pita, C.; Almeida, M. C.; Carreira, I. M.; Gil, A. M. Following Healthy Pregnancy by NMR Metabolomics of Plasma and Correlation to Urine. *Journal of Proteome Research* **2014**, 14, 1263–1274.

Pinto, J.; Almeida, L. M.; Martins, A. S.; Duarte, D.; Domingues, M. R. M.; Barros, A. S.; Galhano, E.; Pita, C.; Almeida, M. C.; Carreira, I. M.; Gil, A. M. Fetal chromosomal disorders impact on maternal metabolome: towards new biomarkers? *American Journal of Obstectrics & Gynecology* **2015**, under revision.

Pinto, J.; Almeida, L. M.; Martins, A. S.; Duarte, D.; Barros, A. S.; Galhano, E.; Pita, C.; Almeida, M. C.; Carreira, I. M.; Gil, A. M. Prediction of gestational diabetes through NMR metabolomics of maternal blood. *Journal of Proteome Research* **2015**, DOI: 10.1021/acs.jproteome.5b00260.

e-Book chapter:

Pinto, J.; Domingues, M. R.; Gil, A. M. Blood Metabolomics in Human Prenatal and Newborn Health Studies. In *Global Metabolic Profiling: Clinical Applications*; Nichols, A., Theodoridis, G., Wilson, I. D., Eds.; Future Science Ltd: London, **2014**; pp 50–68, eISBN: 978-1-909453-61-6.

Contents

Abbreviations and symbols

1,5-AG	1,5-anhydroglucitol
1D	One dimensional
1^{st} T, 2^{nd} T, 3^{rd} T	1^{st}, 2^{nd} and 3^{rd} trimesters
2D	Two dimensional
2Q, 3Q, 4Q, 5Q	Double, triple, quadruple and quintuple quantum
4-DEA	4-deoxyerythronic acid
4-DTA	4-deoxythreonic acid
Ac	Acetate
ACN	Acetonitrile
AFP	Alpha-fetoprotein
Alb	Albumin-lysyl
ANOVA	Analysis of variance
AUC	Area under the curve
BALF	Bronchoalveolar lavage fluid
Bet	Betaine
BCAA	Branched-chain amino acids
BHMT	Betaine homocystein methyltransferase
b_{cvSE}	b coefficient's standard error obtained by cross validation
BMI	Body mass index
BMRB	Biological magnetic resonance bank
Br	Broad
CD	Chromosomal disorders
CICECO	Centro de Investigação em Materiais Cerâmicos e Compósitos
Cit	Citrate
cffDNA	Cell-free fetal DNA
CHD	Congenital heart defects
CNS	Central nervous system
COSY	Correlation spectroscopy
CPMG	Carr-Purcell-Meiboom-Gill
CR	Classification rate
Cr	Creatine
Crn	Creatinine
CSF	Cerebrospinal fluid
CV	Coefficient of variation
CVS	Chorionic villus sampling
d	Doublet
dd	Doublet of doublets
DG	Diacylglycerols
DMA	Dimethylamine
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DOSY	Diffusion ordered spectroscopy
dt	Doublet of triplets
EC	Esterified cholesterol
EDTA	Ethylene diamine tetraacetic acid
EI	Electron impact
ERETIC	Electronic reference to access in vivo concentrations

E.S.	Effect size
ESI	Electrospray ionization
EUROCAT	European surveillance of congenital anomalies
F1	1st dimension
F2	2nd dimension
FA	Fatty acids
FC	Free cholesterol
FID	Free Induction Decay
FM	Fetal malformations
FMOs	Flavin-containing monooxygenases
FN	False negative
FP	False positive
FPG	Fasting plasma glucose
FPR	False positive rate
FT	Fourier transformation
GAA	Guanidinoacetate
GC	Gas chromatography
GDM	Gestational diabetes mellitus
GFR	Glomerular filtration rate
Glc	Glucose
Glyc	Glyceryl
g.w.	Gestational weeks
Hcy	Homocysteine
HDL	High density lipoproteins
HELLP	Hemolysis, elevated liver enzymes and low platelets
HIE	Hypoxic ischemic encephalopathy
HILIC	Hydrophilic interaction chromatography
HMDB	Human metabolome database
HPLC	High performance liquid chromatography
HSQC	Heteronuclear single quantum coherence
HybridSPE-PL	Hybrid solid phase extraction for phospholipids
IEM	Inborn errors of metabolism
INEPT	Insensitive nuclei enhancement by polarization transfer
IR	Infrared spectroscopy
IUGR	Intrauterine growth restriction
J-res	*J*-resolved
LB	Line broadening
LBW	Low birth weight
LC	Liquid chromatography
LC-HRMS	Liquid chromatography high-resolution mass spectrometry
LDL	Low density lipoproteins
LGA	Large for gestational age
LPC	Lysophosphatidylcholines
LPE	Lysophosphatidylethanolamines
LPS	Lysophosphatidylserines
LV	Latent variable
M	Multiplet
MALDI	Matrix-assisted laser desorption and ionization

MBB	Maternity Bissaya Barreto
MeOH	Methanol
MIR	Mid-infrared spectroscopy
MCCV	Monte-Carlo cross validation
MLEV-17	Malcolm Levitt-17 sequence
m/z	Mass-to-charge ratio
n.d.	Not determined
MQ	Multiple quantum spectroscopy
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MTBE	Methyl tert-butyl ether
MVA	Multivariate analysis
M_w	Molecular weight
NB	Newborn
NMR	Nuclear magnetic resonances spectroscopy
NP	Non-pregnant
NS	Number of scans
NTD	Neural tube defects
OGTT	Oral glucose tolerance test
OPLS-DA	Orthogonal partial least squares-discriminant analysis
p	p-value
PA	Phosphatidic acids
PAPP-A	Pregnancy-associated plasma protein A
PC	Phosphatidylcholines
PCA	Principal Component Analysis
PE	Phosphatidylethanolamines
PG	Phosphatidylglycerols
PI	Phosphatidylinositols
PL	Phospholipids
PLS	Partial least squares regression
PLS-DA	Partial least squares-discriminant analysis
PQN	Probababilistic quotient normalization
PROM	Premature rupture of membranes
PS	Phosphatidylserines
PTD	Preterm delivery
Pyr	Pyruvate
q	Quartet
Q^2	Goodness of prediction or prediction power
Q-IT	Quadrupole ion trap
r	Correlation coefficient
R^2_x	Variation explained by the **X** matrix
R^2_y	Variation explained by the **Y** matrix
RDS	Respiratory distress syndrome
RF	Radiofrequency
RG	Receiver gain
ROC	Receiver operating characteristic
RSPA	Recursive segment-wise peak alignment
RT	Retention time

s	singlet
Sens	Sensitivity
SGA	Small for gestational age
SI	Size of real spectrum
SIM	Selected ion monitoring
SM	Sphingomyelins
SOPs	Standard operating procedures
SPE	Solid phase extraction
Spec	Specificity
SR	Spectrum reference frequency
SRM	Selected reaction monitoring
STOCSY	Statistical total correlation spectroscopy
SW	Spectral window
t	Triplet
T13, T18, T21	Trisomies 13, 18 and 21
TCA	Tricarboxilic acid
TD	Size of FID
TG	Triacylglycerides
TIC	Total ion chromatogram
TMA	Trimethylamine
TMS	Tetramethylsilane
TMAO	Trimethylamine N-oxide
TN	True negative
TOF	Time of flight
TOCSY	Total correlation spectroscopy
TP	True positive
TPR	True positive rate
TSP	3-trimethylsilylpropionic acid
UA	University of Aveiro
Un_i	Unassigned resonance i
uE3	Unconjugated estriol
UPLC	Ultraperformance liquid chromatography
UV	Unit variance
VIP	Variable importance to the projection
VIP_{cvSE}	VIP's standard error obtained by cross validation
VLBW	Very low birth weight
VLDL	Very low density lipoproteins
WDW	Type of window function
β-hCG	Free β-human chorionic gonadotrophin
β-HB	β-hydroxybutyrate

Chapter 1. Introduction

1.1. Pregnancy and fetal development

A human full-term pregnancy can range from 37 to 42 gestational weeks (g.w.), counting from the first day of the last normal menstrual period (Creasy et al., 2009). Gestational weeks are grouped into three trimesters, each one characterized by specific hallmarks of fetal development. The first trimester (1st T) corresponds to the period from fertilization until 14 g.w., during which the most important steps of fetal organogenesis occur constituting the period of greatest risk for birth defects. At this stage, important fetal-maternal structures become completely differentiated (e.g. chorionic villi and capillary system which form the placenta and umbilical cord, respectively) forming the definitive annexes that support fetal growth as well as maintenance, regulatory and endocrine functions. The following trimester (2nd T) is defined by the period from 14 to 28 g.w., mainly characterized by organ maturation and fetal growth. The third trimester starts on the 28th week and extends to the end of pregnancy and it is during this time that fetal growth is enhanced and the fetus becomes able to synthesize its own body fat. By the 37th week, fetal maturation is complete and pregnancy is at term, although labour can occur up to 41 or 42 weeks.

During the time course of pregnancy, important maternal-placental-fetal metabolic interactions occur in order to 1) ensure adequate growth and development of the fetus, 2) provide the fetus with adequate energy stores and substrates needed after birth, 3) ensure maternal needs for the increased physiological demands of pregnancy and 4) provide the mother with sufficient energy stores and substrates required for labour and lactation (Hadden and McLaughlin, 2009). The transfer of nutrients and other substances between fetal and maternal circulations, in both directions, is illustrated in Figure 1.1. The placenta is the primary site of nutrient and gas exchange between the mother and fetus presenting three main functions: ensuring fetal metabolism (e.g. synthesis of glycogen), transport of gases and nutrients and endocrine excretion. Only selected nutrients from the maternal blood are able to cross the placenta, glucose being the main nutrient transported and a primary source of energy to the fetus. Glucose is responsible for maintenance of basal metabolism, energy storage in form of glycogen and in adipose tissues, as well as providing energy for protein synthesis and growth. In addition, the active transfer of amino acids from maternal circulation ensures the structural basis for protein synthesis, essential for fetal growth, and provides the oxidative substrates for energy production, mainly when glucose levels are low. Lipids cross the placenta with difficulty, although essential fatty acids and long-chain polyunsaturated fatty acids cross from maternal circulation to fetus (Herrera and

Ortega, 2008; Moore and Persaud, 2008). Harmful substances also have the ability to cross the placenta comprising drugs, chemicals and infection agents.

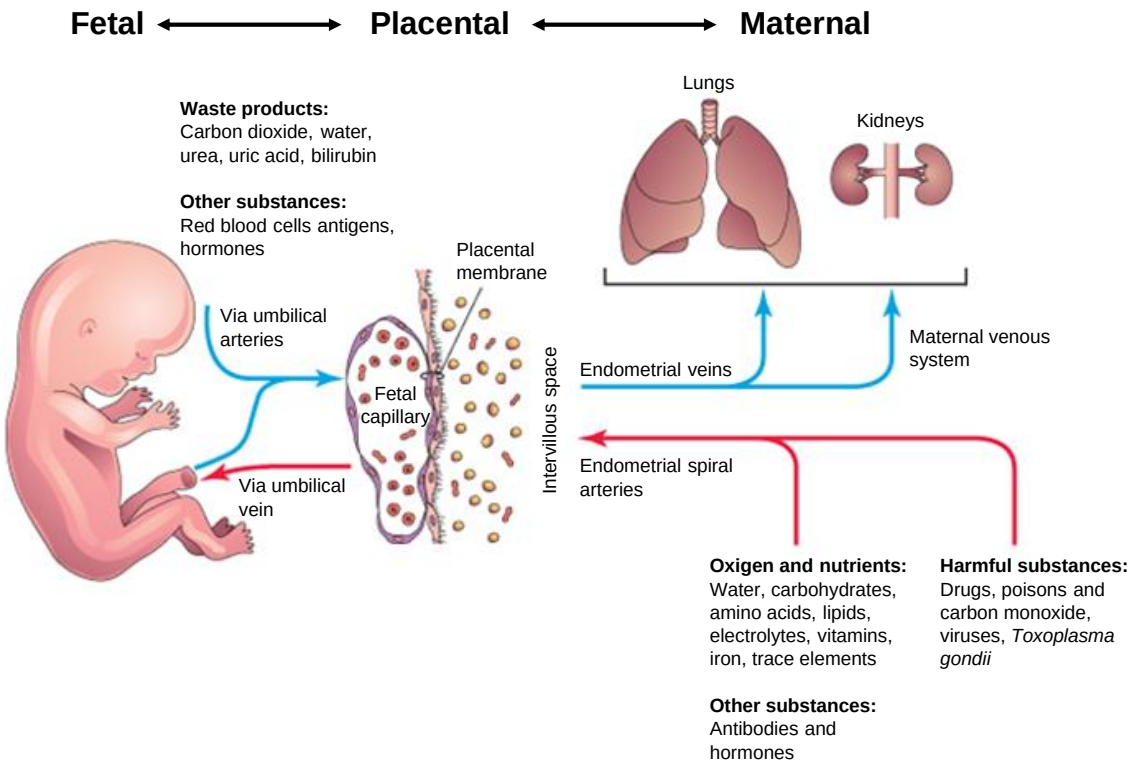


Figure 1.1. Transfer processes taking place between maternal, placental and fetal environments. [Adapted from (Moore and Persaud, 2008)]

The transfer of nutrients from mother to fetus leads to several metabolic adaptations that take place in maternal metabolism to satisfy both maternal and fetal needs. These metabolic changes include hyperinsulinemia, insulin resistance, relative fasting hypoglycemia, increased circulating plasma lipids and hypoaminoacidemia (Creasy et al., 2009). Although the precise mechanism of insulin resistance is not entirely clear, hormonal alterations throughout gestation have been attributed to the reduced insulin sensitivity. Hypoglycemia is a result of the high rate of placental transfer of glucose, despite of enhanced gluconeogenesis and reduced consumption of glucose. Total plasma lipids increase significantly and progressively after 24 g.w., with increases in triglycerides, cholesterol and free fatty acids being most marked. Hypoaminoacidemia occurs due the above mentioned increased placental uptake of amino acids leading to diminished maternal circulating levels (Butte, 2000; Herrera and Ortega, 2008; Creasy et al., 2009).

1.2. Prenatal disorders and diagnostic procedures: present status and challenges

Despite all metabolic and physiological alterations occurring throughout gestation, disruptions may occur and cause maternal and fetal health complications. Prenatal disorders comprise maternal and fetal disorders, although most of them affect both mother and fetus to some extent. The main pathologies are listed in Table 1.1, in decreasing order of prevalence.

The work described in this thesis entailed the study of prenatal disorders with higher prevalence, namely maternal disorders such as preterm delivery (PTD), gestational diabetes mellitus (GDM), preeclampsia, and premature rupture of the membranes (PROM), and fetal disorders such as chromosomal disorders (CD) and fetal malformations (FM). For other prenatal disorders such as intrauterine growth restriction (IUGR), small for gestational age (SGA) and large for gestational age (LGA), it was not possible to gather sufficient number of samples. A brief overview (i.e., pathophysiology, diagnostic procedures and management) of each of these conditions is presented throughout this section.

Table 1.1. Most common maternal and fetal disorders with corresponding prevalence.

Prenatal disorder	Prevalence	Reference
Maternal		
Preterm delivery (PTD)	11.1% of all live births	(Blencowe et al., 2013)
Gestational Diabetes Mellitus (GDM)	9.2% of all pregnancies	(DeSisto et al., 2014)
Preeclampsia	3-8% of all pregnancies	(Uzan J et al., 2011)
Premature rupture of the membranes (PROM)	8% of all pregnancies	(Hannah et al., 1996)
Fetal		
Intrauterine growth restriction (IUGR)	7-15% of all pregnancies	(Cetin and Alvino, 2009)
Small for gestational age (SGA)	<10% of all live births	(Horgan et al., 2010)
Chromosomal disorders (CD)	3-4% of all pregnancies	(Creasy et al., 2009)
Fetal malformations (FM)	2-3% of all live births	(Creasy et al., 2009)

1.2.1. Preterm delivery

Preterm delivery (PTD), also referred to as preterm birth, is defined by the World Health Organization as any births occurring before 37 completed weeks of gestation or fewer than 259 days since the first day of a woman's last menstrual period (Blencowe et al., 2013). PTD occurs in 11.1% of all live births and is the major cause of neonatal mortality and morbidity. Long-term adverse consequences for children who are born prematurely include cerebral palsy, sensory deficits, learning disabilities and respiratory

illnesses compared with children born at term (Beck et al., 2010; Blencowe et al., 2013). PTD can be further subdivided based on gestational age: extremely preterm (<28 g.w.), very preterm (28 - <32 g.w.) and moderate preterm (32 - <37 completed g.w.). Lower gestational age at birth has been associated with greater risks. (Blencowe et al., 2013).

The causes of PTD can be classified into two broad subtypes: 1) spontaneous PTD defined as spontaneous onset of labour or following prelabour premature rupture of membranes (pPROM), and 2) provider-initiated preterm delivery defined as induction of labour or elective caesarean birth before 37 completed g.w. for maternal or fetal indication (e.g., placenta accrete) (Blencowe et al., 2013). Risk factors for spontaneous PTD are higher maternal age, multiple pregnancy, infection, underlying maternal chronic medical conditions, nutritional, lifestyle/work related, maternal psychological health, genetic among others.

PTD diagnosis remains a challenge due to the incompletely understood sequence and timing of events that precede preterm labour. Standard criteria for diagnosis is based on the measurement of uterine contractions (>4 per 20 minutes or 8 per 60 minutes) accompanied by cervical change (cervical effacement of 80% or cervical dilatation of 2 cm or more), but it lacks precision. The result is a false-positive diagnosis in 40% of women and enrolment of women who are not in labour into trials of agents to arrest PTD (Hacker et al., 2004; Creasy et al., 2009). Due the high rates of neonatal morbidity and mortality and the later in life health complications of children born preterm, early detection preceding the onset of PTD is a recognised requirement. The identification of high risk women might allow the initiation of specific pre-treatment care and provide important insight into the mechanisms leading to preterm birth.

1.2.2. Gestational diabetes mellitus

Gestational diabetes mellitus (GDM) is defined as “a carbohydrate intolerance of varying degrees of severity with onset or first recognition during pregnancy” and is usually diagnosed in the 2nd T between 24 and 28 weeks of gestation (Creasy et al., 2009). In most instances, the underlying pathophysiology of GDM is similar to that observed in type 2 diabetes characterized by an inability to maintain an adequate insulin response due the significant decreases in insulin sensitivity with advanced gestation. This condition often occurs in subjects with no history of diabetes, and factors such as age (older than 35 to 40 years), obesity (body mass index (BMI) >30), history of previous GDM or large for gestational age (LGA) infant, and genetic

background contribute to the severity of the disease (Creasy et al., 2009). Clinically recognition of GDM is important since therapy can reduce pregnancy complications and potentially reduce long-term sequelae in the offspring. General literature have documented an increased risk of GDM mothers and babies for the development of other prenatal complications such as preeclampsia, metabolic syndrome, birth trauma, fetal hyperinsulinemia, macrosomia, stillbirth, newborn hypoglycemia and hyperbilirubinemia, among others (Creasy et al., 2009; Pridjian and Benjamin, 2010).

Currently, GDM diagnosis involves two discrete temporal phases: a) measurement of fasting plasma glucose (FPG) performed at the first prenatal visit and b) oral glucose tolerance test (OGTT) performed at 24-28 g.w. of gestation (Metzger et al., 2010). In the first prenatal visit, FPG results are interpreted according three thresholds: 1) if FPG is ≥126 mg/dL (7.0 mmol/L), treatment and follow-up as for pre-existing diabetes is applied, 2) if FPG ≥92 mg/dL (5.1 mmol/L) but <126 mg/dL (7.0 mmol/L), the GDM diagnosis is made, and 3) if FPG <92 mg/dL (5.1 mmol/L), a new test for GDM is performed at 24-28 g.w. with a 75 g OGTT. The GDM diagnosis at 24-28 g.w. is established if one or more values of the 75 g OGTT equals or exceeds the thresholds indicated in Table 1.2 (Metzger et al., 2010; George, 2011).

Table 1.2. Threshold values for diagnosis of GDM. [Adapted from (Metzger et al., 2010; George, 2011)]

Glucose measurement	**75 g OGTT**	
	mg/dL	**mmol/L**
Fasting	≥92	≥5.1
1 hour	≥180	≥10.0
2 hours	≥153	≥8.5

GDM management includes dietary therapy, exercise, and medical treatment (Pridjian and Benjamin, 2010). In dietary therapy, GDM patients are advised to have five meals a day and the daily calorie requirement is calculated according to the standard of 1900-2400 kcal per day with carbohydrate restriction to 35-40% of calories. If dietary control is carried out strictly, most patients do not need medical treatment like insulin therapy. Physical exercise may increase the sensitivity of peripheral tissues to insulin, thus being able to help to control glucose level of GDM patients.

1.2.3. Preeclampsia

Preeclampsia is a disorder that occurs only during pregnancy, characterized by the new onset of hypertension, proteinuria and generalized edema, usually after the 20th week of gestation, with an unpredictable course that affects both mother and fetus (Creasy et al., 2009). There are numerous risk factors for development of preeclampsia, including genetic factors, nulliparity, a new partner, demographic factors (e.g. maternal age >35 years), factors related to the pregnancy (e.g., multiple pregnancy, congenital or chromosome anomalies, urinary infection), factors associated with maternal disease (e.g. chronic hypertension, kidney disease, obesity, insulin resistance, diabetes), and environmental factors (e.g., living at a high altitude and stress) (Uzan J et al., 2011). The placenta plays a key role in the underlying mechanisms leading to the development of preeclampsia including inadequate placental cytotrophoblast invasion, followed by widespread maternal endothelial dysfunction. In addition, excessive oxidative stress, inflammation, immune maladaptation, and genetic susceptibility may all contribute to the pathogenesis of preeclampsia (Young et al., 2010).

Preeclampsia has been characterized by some researchers into two different disease entities: early-onset and late-onset preeclampsia. Early-onset preeclampsia is usually defined as preeclampsia that develops before 34 weeks of gestation, and has severe maternal and fetal consequences including IUGR, PTD, low- or very-low birth weight (LBW or VLBW), increased perinatal morbidity and mortality and a high incidence of the life-threatening HELLP syndrome (hemolysis, elevated liver enzymes, and low platelets) (Raymond and Peterson, 2011; Than et al., 2012). On the other hand, late-onset preeclampsia develops at or after 34 weeks of gestation and its clinical presentation is frequently mild, resulting mainly in maternal consequences.

The diagnosis of preeclampia is defined by an increase in maternal blood pressure and high levels of protein in urine (proteinuria) with onset beyond the 20th week of gestation (Creasy et al., 2009; Uzan J et al., 2011). The threshold values for diagnosis are 24 hour proteinuria ≥300 mg/day, and a systolic blood pressure >140 mmHg or diastolic blood pressure ≥90 mmHg. When the diagnosis is made, the only cure still known is immediate delivery, which results in high number of babies born prematurely. Accurate prediction of preeclampsia would enable early and optimal management of women at high risk. Several predictive markers have been investigated, such as maternal mean arterial pressure, uterine artery Doppler pulsatility index, and 1st T maternal serum markers (e.g., placental growth factor, pregnancy-associated plasma protein A (PAPP-A) and free β-human chorionic gonadotrophin (β-hCG)), but these

lack reliable detection and have high false-positive rates (80.8-93% detection rates with 5-10% false-positive rates) (Kane et al., 2014).

1.2.4. Premature rupture of membranes

Premature rupture of the membranes (PROM) is defined as the spontaneous rupture of membranes (or amniotic sac) at term before the onset of contractions. Rupture of fetal membranes is an integral part of the normal parturition process at term and usually occurs as a result of a physiologic process of progressive membrane weakening (Creasy et al., 2009). PROM affects approximately 8% of pregnancies at term, and 95% of these women deliver within 28 hours of membrane rupture. The expeditious delivery of the PROM patient can reduce the risk of perinatal infections without increasing the likelihood of operative delivery (Hannah et al., 1996; Creasy et al., 2009).

PROM diagnosis is based on the history of vaginal loss of fluid and confirmation of amniotic fluid in the vagina. PROM diagnosis is made by 1) testing of the fluid with Nitrazine paper, which will turn in blue in the presence of alkaline amniotic fluid (pH of 7.1 to 7.3), and 2) microscopic inspection for the presence of arborized crystals (i.e., ferning, an indication of the presence of estrogen) in an air-dried sample collected from the vaginal side walls or pooled vaginal fluid. False-positive nitrazine test results occur in the presence of alkaline urine, blood, or cervical mucus (Hacker et al., 2004; Creasy et al., 2009).

1.2.5. Fetal growth anomalies

Intrauterine growth restriction (IUGR) and small for gestational age (SGA) are fetal disorders with high occurrence in pregnancy, 7-15% and <10%, respectively (Cetin and Alvino, 2009; Horgan et al., 2010). Although some overlap occurs among these two disorders, they are differently defined. IUGR is defined as a fetus that does not reach its growth potential and is characterized at birth by a weight or a BMI below normal values for the corresponding number of gestational weeks. On the other hand, SGA defines an infant that has a birth weight below the 10^{th} percentile for specific gestational age. It is important to note that not all SGA infants identified at birth are patients with IUGR, and that, on the other hand, infants with IUGR are not necessarily classified as SGA (Dessì et al., 2013). There are several factors contributing to these disorders such as genetic factors, disturbances in the ability of the mother to provide sufficient nutrients or in the ability of the placenta to transfer nutrients, intrauterine

hormones, and growth factors (Pallotto and Kilbride, 2006). Both conditions are known as key risk factors for increased perinatal mortality, birth adaptation complications, including perinatal acidosis, hypoglycaemia, hypothermia, coagulation abnormalities, and selected immunologic deficiencies. In addition, IUGR infants also have greater risk for complications of prematurity (e.g., chronic lung disease) and increased risk for short stature, cognitive delays and neurologic disorders in childhood (Pallotto and Kilbride, 2006). In contrast, large for gestational age (LGA) or macrosomia is defined as a neonate with a birth weight above the 90th centile for gestational age. The major risk factors for macrosomia are maternal diabetes, obesity and excessive nutrition/weight gain during pregnancy. The LGA baby often undergoes hypoglycaemia caused by chronic fetal hyperglycaemia which continues even after birth (Dessì et al., 2013).

Abnormal fetal growth is detected with the clinical suspicion of a subnormal uterine size and the diagnosis is usually made by ultrasound. However, there is a high percentage of error in this and fetal growth complications are undetected in about 30% of routinely scanned cases and incorrectly detected in 50% of cases (Bamfo and Odibo, 2011).

1.2.6. Chromosomal disorders

Chromosomal disorders (CD) are defined as changes resulting in a visible alteration of the chromosomes produced by specific chromosomal mechanisms (Moore and Persaud, 2008). Figure 1.2 shows an example of the normal human karyotype comprising 22 pairs of chromosomes (autosomes), and a pair of sex chromosomes (XX in females and XY in males). The common chromosome aberrations in humans are abnormalities in chromosome number and alterations of chromosome structure (Moore and Persaud, 2008; Creasy et al., 2009). Table 1.3 shows the main characteristics of each type of chromosomal disorder along with some examples. Numerical abnormalities represent either aneuploidy or polyploidy. Aneuploidy occurs during both meiosis and mitosis and is characterized by a number of chromosomes that is not an exact multiple of the haploid number of 23 (e.g., 45 or 47). Polyploidy is characterized by a chromosome number that is a multiple of the haploid number of 23 other than the diploid number (e.g., 69). Aneuploidy is the most frequently seen chromosome abnormality in clinical cytogenetics (3% to 4% of clinically recognized pregnancies), including the autosomal trisomies, as well as sex chromosomal aneuploidies (Table 1.3). Trisomy 21 (T21), trisomy 18 (T18) and Trisomy 13 (T13) are the most common

autossomal numerical disorders in liveborn humans with an approximate incidence of 1/830, 1/7500 and 1/22700, respectively (Moore and Persaud, 2008).

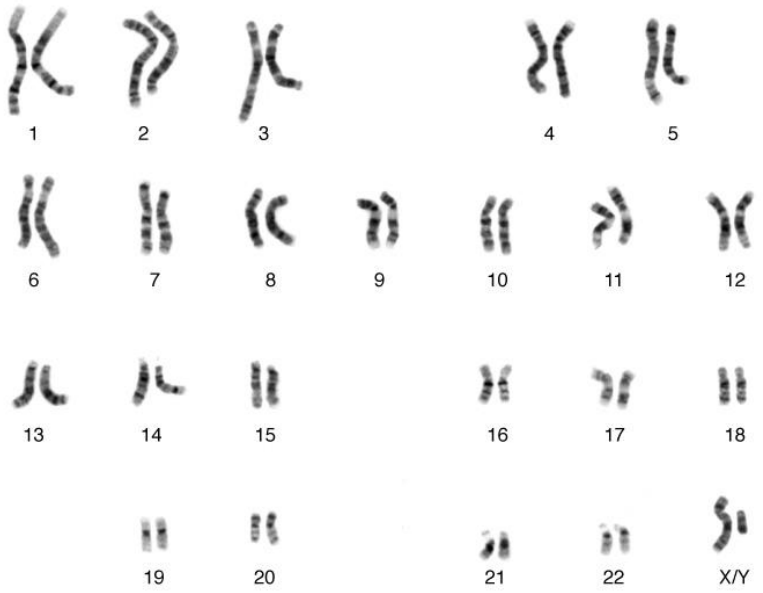


Figure 1.2. The human karyotype comprising 22 pairs of chromosomes and a pair of sex chromosomes (XX in females and XY in males). Reproduced from "Talking Glossary of Genetic Terms" (http://www.genome.gov/glossary/index.cfm?id=114).

Most of the structural chromosomal abnormalities result from chromosome breakage followed by reconstitution in an abnormal combination (Moore and Persaud, 2008; Creasy et al., 2009). Chromosome breakage may be induced by various environment factors such as radiation, drugs, chemicals and viruses. Structural abnormalities are balanced if there is no net loss or gain of chromosomal material, or unbalanced if there is an abnormal genetic complement. The resulting type of structural chromosomal abnormality depends on what happens to the broken pieces and can be subdivided into deletions, duplications, insertions, inversions, translocations and isochromosomes as described in Table 1.3.

Depending on the type, chromosome abnormalities may range from small phenotypic alterations to more severe developmental impairment such as mental deficiency and malformations. For instance, the phenotypic changes of individuals born with T21 are mental deficiency, brachycephaly, flat nasal bridge, upward slant to palpebral fissues, protruding tongue, simian crease, clinodactyly of fifth digit, congenital heart defects and gastrointestinal tract anomalies (Moore and Persaud, 2008).

Table 1.3. Some of the most common chromosomal abnormalities. [Adapted from (Moore and Persaud, 2008; Creasy et al., 2009)]

Abnormality	Characteristics
Numerical	**Aneuploidy:** • **Monosomy** – presence of only one chromosome from a pair Example: Turner syndrome (45, X) • **Trisomy** – presence of an extra third chromosome Examples: Trisomy 21 or Down syndrome (47,XX+21 or 47,XY+21), Trisomy 18 or Edwards syndrome (47,XX+18 or 47,XY+18), Trisomy 13 or Patau syndrome (47,XX+13 or 47,XY+13), Kleinfelter syndrome (47,XXY), Triple X syndrome (47,XXX) and 47,XYY syndrome **Polyploidy:** Example: 48,XXXX, 48,XXXY, 48,XXYY and 49,XXXXY
Structural	**Deletions:** loss of a chromosome segment **Duplications:** duplication of chromosome segments **Insertions:** insertion of a chromosome segment into a nonhomologous chromosome **Inversions:** a segment of a chromosome is reversed **Translocations:** transfer of chromosomal material between chromosomes • **Reciprocal translocation** – exchange of segments from two different chromosomes • **Robertsonian translocation** – attachment of an entire chromosome to another at the centromere **Isochromosomes:** one arm of a chromosome is lost and the other arm is duplicated

Non-invasive screening tests for fetal chromosomal disorders may be offered in the 1st and 2nd trimesters by ultrasound and maternal serum concentrations of various fetoplacental products, such as free β-hCG, PAPP-A, alpha-fetoprotein (AFP), unconjugated estriol (uE3) and inhibin A (Nicolaides, 2011). In Table 1.4, common screening and diagnostic tests are listed including timing of procedures, detection rate and false-positive rates. Earlier effective screening for chromosomal abnormalities may be provided at 11-14 g.w. by a combination of maternal age, fetal nuchal translucency thickness and maternal serum markers including free β-hCG and PAPP-A. This method allows the early screening of T21 and other aneuploidies with a detection rate of 85-95% with 5% false-positive rate, and >90% with 1% false-positive rate, respectively (Table 1.4) (Nicolaides, 2004, 2011). In the 2nd T, different combinations of maternal age with maternal serum markers can be offered for screening for aneuploidies. As shown in Table 1.4, the quadruple test based on the combination of maternal age with serum AFP, free β-hCG, uE3 and inhibin A provides a higher detection rate than double (maternal age with serum AFP, free β-hCG) and triple (maternal age with serum

AFP, free β-hCG, uE3) tests for T21 (Nicolaides, 2011). Detection of chromosomal abnormalities by ultrasound alone will depend on whether there are associated physical malformations. For instance, the detection rate for T21 in 2nd T have been 79.9% with a false-positive rate of 6.7% (Benn et al., 2002; Collins and Impey, 2012).

Table 1.4. Common prenatal screening and diagnostic procedures for chromosomal disorders. [a]Combined test: combination of maternal age, nuchal translucency, free β-hCG and PAPP-A; [b]Trisomies 18 and 13, Turner syndrome, and triploidy; [c]Double test: combination of maternal age with serum AFP and free β-hCG; [d]Triple test: combination of maternal age with serum AFP, free β-hCG and uE_3; [e]Quadruple test: combination of maternal age with serum AFP, free β-hCG, uE_3 and inhibin A. T21: Trisomy 21

Procedure (time for results)	Timing of procedures	Detection /false-positive rates	References
Screening			
1st T maternal serum markers (2-3 days)	11-14 g.w.	Combined test[a]: T21: 85-95% /5% Others[b]: >90% /1%	(Nicolaides, 2004, 2011)
2nd T maternal serum markers (2-3 days)	15-20 g.w.	Double test[c]: T21: 60-65% /5% Triple test[d]: T21: 65-70% /5% Quadruple test[e]: T21: 70-75% /5%	(Nicolaides, 2011)
Diagnosis			
Ultrasound (immediately)	11-14 g.w. & 18-22 g.w.	T21: 79.9% /6.7% (at 18-22 g.w.)	(Benn et al., 2002)
Chorionic villus sampling (CVS) (up to 2 weeks)	10-13 g.w.	T21 and others[b]: 97.5–99.6%	(Hahnemann and Vejerslev, 1997)
Amniocentesis (up to 2 weeks)	15-20 g.w.	T21 and others[b]: 99.4–99.8%	(Hahnemann and Vejerslev, 1997)
Cordocentesis (3 days)	>18 g.w.	-	(Moore and Persaud, 2008)

Invasive prenatal diagnosis is offered to women who screen above a set risk cut-off level on non-invasive screening or to pregnant women who's personal, obstetrical, or family history places them at increased risk of having an offspring with chromosomal abnormalities (Creasy et al., 2009). The main invasive diagnostic tests used for chromosomal anomalies detection are chorionic villus sampling (CVS, performed in the 1st T), amniocentesis (performed in 2nd T) or cordocentesis (percutaneous umbilical blood sampling performed after 18 g.w.) (Keeling and Khong, 2007). These methods are very invasive representing a percentage of risk for mother and fetus due the associated complications including vaginal spotting, amniotic fluid leakage, fetal needle injury and fetal loss. The fetal loss rate associated with amniocentesis and CVS is often reported to be 0.5-1%, while for cordocentesis is 3.2% (Tongsong et al., 2000; Tabor and Alfirevic, 2010). Another disadvantage is the typically high length of time required

for the results of amniocentesis and CVS diagnostic procedures to become available (1-2 weeks).

Recently, analysis of cell-free fetal DNA (cffDNA) in maternal blood has been shown to have higher detection rates than the currently available screening tests for fetal autosomal aneuploidies. Results of studies using 1st T cffDNA have been shown to detection rates of 98.6-100%, 100% and 91.7% for T21, T18 and T13, respectively, with false-positive rates < 1% (Walsh and Goldberg, 2013). However, cffDNA cannot currently replace either CVS or amniocentesis due the lower detection rate obtained for T13 and high costs (ranging from about $500 to $1500 per test). In addition, karyotyping can provide information about other conditions (e.g., open neural tube defects) besides fetal aneuploidy, whereas cffDNA can not (Benn et al., 2013; Walsh and Goldberg, 2013). Despite the clinical usefulness of these screening and diagnostic procedures, the development of reliable tests with higher detection rate and lower false-positive rates is still required. In addition, the development of cost-effective and easy to perform genetic screening tests would extend to the less-developed countries.

1.2.7. Fetal malformations

Fetal malformations (FM), also referred as congenital anatomic anomalies, congenital malformations or birth defects, are defined as a structural abnormality of any type that is present at birth (Moore and Persaud, 2008). FM may be single or multiple and of minor or major clinical significance. Usually, abnormalities of the organ systems occur during fetal development and these comprise neural tube defects, ventricular septal defects, cleft lip, cleft palate, low-set malformed ears, and deafness, among others. Some of these structural anomalies occur early in the main embryonic period (3-8 g.w) when organ systems are formed, but also later during the growth and maturation of the fetus (9-38 g.w.), when organs such as the cerebral cortex, the gastrointestinal tract and the renal glomeruli continue to differentiate and develop (Moore and Persaud, 2008). Table 1.5 lists the major fetal malformations occurring in human organs at birth with corresponding prevalence per 10 000 births in Europe (EUROCAT Working Group, 2012). The European surveillance of congenital anomalies (EUROCAT) reported highest prevalence for congenital anomalies of cardiovascular system from 2008 to 2012, followed by limb and urinary systems. Within cardiovascular system defects, the most common have been ventricular septal defect (31.0 per 10 000 live births) and atrial septal defect (19.4 per 10 000 live births). Some anomalies (e.g. renal agenesis, anencephaly) are always lethal, with death occurring either in utero or in the early neonatal period. Other anomalies are compatible with limited survival but

associated with very severe neurologic disabilities (e.g., hydrencephaly), for which early pregnancy termination is the usual management. For most anomalies, the prognosis is good but uncertain (e.g., mild severe ventriculomegaly) with outcomes ranging from severe physical limitation (e.g., spina bifida) or uncertain mental faculties. (e.g., agenesis of the corpus callosum) (Creasy et al., 2009).

Table 1.5. Major malformations occurring in human organs at birth with corresponding prevalence per 10 000 live births in Europe from 2008 to 2012. [Adapted from (EUROCAT Working Group, 2012)]

Organ system	Prevalence per 10 000 live births	Types of fetal malformations
Central nervous system (CNS)	22.2	Neural tube defects, spina bifida, encephalocele, anencephaly, hydrocephaly, microcephaly, holoprosencephaly
Eye	3.6	Anophtalmia/microphtalmia, congenital cataract, congenital glaucoma
Cardiovascular	71.9	Common arterial truncus, transposition of great vessels, ventricular septal defect, atrial septal defect, tetralogy of Fallop, hypoplastic left heart, pulmonary valve stenosis, pulmonary valve atresia
Respiratory	3.8	Choanal atresia, cystic adenomatous malformation of lung
Oro-facial	12.7	Cleft palate, cleft lip
Digestive	16.6	Oesophageal atresia, duodenal atresia or stenosis, ano-rectal atresia and stenosis, diaphragmatic hernia, abdominal wall defects, gastroschisis, omphalocele
Urinary	32.7	Multicystic renal dysplasia, congenital hydronephrosis, bilateral renal agenesis
Genital	21.6	Hypospadias
Limb	37.9	Limb reduction defects, club foot, hip dislocation and/or dysplasia, polydactyly, syndactyly

Currently, two-dimensional ultrasound, performed in the 1st (11-14 g.w.) and 2nd (18-22 g.w.) T, is the primary method for detection of fetal structural defects, and other methods such as amniocentesis (15-20 g.w.), are reserved for specific or confirmatory circumstances (Creasy et al., 2009). The sensitivity of ultrasound for detection of major anomalies is variable but in general surprisingly low, with reported detection rates from 22% to 55%. Ultrasound detection rates are dependent on fetal anomaly type, gestational age, risk factors and imaging method and quality. The best detection rates by ultrasound are obtained for major CNS anomalies such as anencephaly (94-100%), with major fetal cardiac defects in the mid range (25-60%) and clefts tending to be lower (Creasy et al., 2009). Amniocentesis is performed in order to measure the alpha-fetoprotein (AFP) concentration in the amniotic fluid. Amniotic fluid AFP concentration

is high in pregnancies whose fetuses have severe anomalies of CNS and ventral abdominal wall, and when combined with ultrasound scanning, enables the diagnosis of approximately 99% of fetuses with these severe anomalies. In addition, when a fetus has an open neural tube defect (NTD), the concentration of AFP in maternal serum is higher than normal due the rapidly diffusion of this protein from fetal tissues into maternal serum via fetal urine (Moore and Persaud, 2008). However, maternal serum AFP measured at 11-13 g.w. has provided detection rates for NTDs ranging from 37.5% to 50% with false positive rates of 10% and 5%, respectively (Bredaki et al., 2012).

1.3. Metabolomics and lipidomics methods applied in disease diagnosis and prognosis

1.3.1. Concept and strategy

Novel approaches have emerged to allow the qualitative and quantitative measurements of complex interactions in biological systems, namely genomics (the study of genes), transcriptomics (the study of gene expression), proteomics (the study of protein expression), metabolomics and lipidomics, referred globally as “omic” technologies (Figure 1.3). Metabolomics, a term today often used interchangeably with the term “metabonomics”, relates to an analytical strategy that entails the comprehensive analysis of all metabolites involved in the metabolic networks of living organisms and their response to pathophysiological or other stimuli (Nicholson et al., 1999). Lipidomics can be viewed as a sub-discipline of metabolomics and is defined as the comprehensive identification and quantification of all lipid molecular species in a biological system (Wenk, 2005; Griffiths and Wang, 2009). It is thus an extension of lipid profiling but put on a quantitative scale and performed in an “omic” manner (i.e., analysis of multiple samples in a high throughput setting).

The general steps performed in metabolomic and lipidomic studies are illustrated in Figure 1.4. Both metabolomics and lipidomics use a variety of biological samples, such as biofluids, tissues and cells, with the aim to investigate their composition and find differences in the nature or amounts of metabolites that characterize phenotypes and biochemical responses to perturbations (e.g., diseases, pharmacological treatment). In metabolomic studies, biological samples are subjected to a minimal sample preparation protocol in order to detect as much metabolites as possible. In lipidomics, a lipid extraction step is usually indispensable before analysis, in order to remove interfering

agents, such as proteins, saccharides, or other small molecules. Several lipid extraction methods have been used, taking advantages of the utilization of solid-liquid or liquid-liquid extractions, such as chloroform/methanol methods (e.g., Folch, Bligh & Dyer), hexane/isopropanol, methyl tert-butyl ether (MTBE)/methanol and solid phase extraction (SPE) (Li et al., 2014).

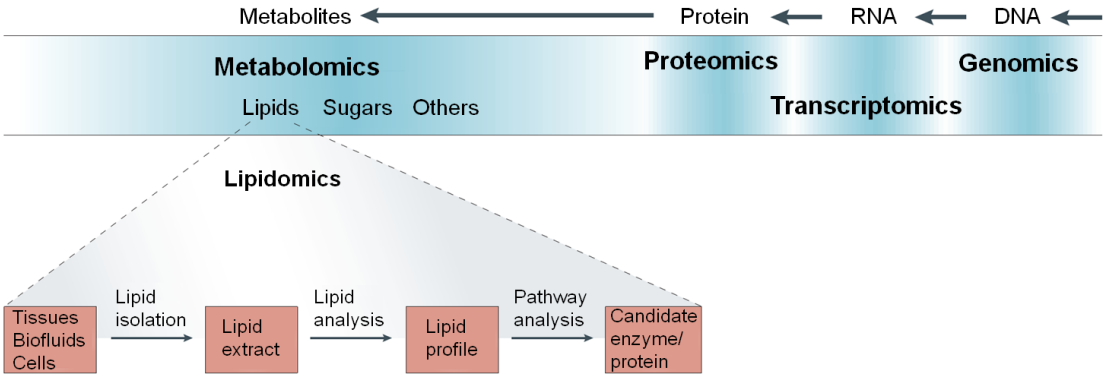


Figure 1.3. Schematic representation of the "Omic" technologies. [Adapted from (Wenk, 2005)]

The analysis of complex biological samples involves the use of advanced metabolite profiling techniques, such as Nuclear Magnetic Resonance (NMR) spectroscopy and Mass spectrometry (MS). While NMR is commonly applied for global metabolic profiling, MS is applied for both global and targeted profiling, although the latter is still a more common approach. Global metabolic profiling (untargeted) studies are focused on the investigation of all the metabolites detectable in a sample capturing as much information as possible. On the other hand, targeted profiling is focused on a small number of metabolites (e.g., phospholipids), commonly requiring a prior isolation step followed by separation using chromatographic methods (gas chromatography (GC) or liquid chromatography (LC)), or using specific and targeted technological approaches such as neutral loss scan or parent ion scan modes in MS analysis. The use of both NMR and MS techniques is often necessary for full molecular characterization in global metabolite profiling (Collino et al., 2009). On the other hand, lipidomics analysis is mainly focused in MS techniques (mainly LC-MS). The benefits, drawbacks and costs of the NMR and MS techniques are summarized in Table 1.6. Other analytical techniques less commonly used in metabolomics are Raman and Infrared spectroscopies (Ellis and Goodacre, 2006).

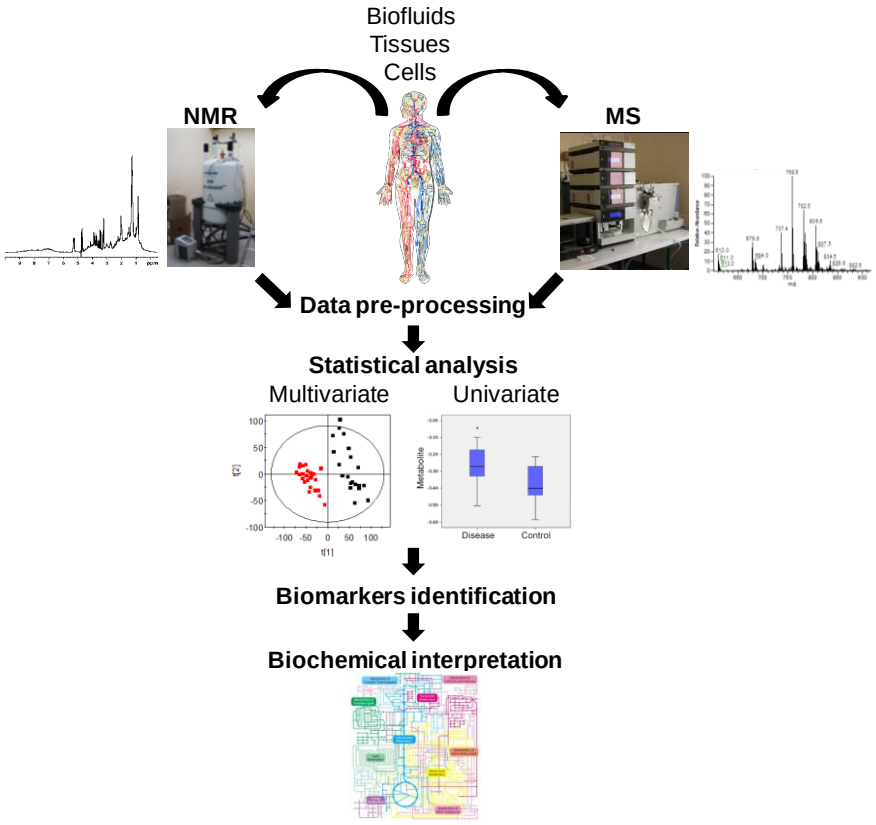


Figure 1.4. Illustrative scheme of a typical metabolomics flowchart.

Table 1.6. Comparison between the analytical techniques commonly used in metabolomics. NMR: nuclear magnetic resonance spectroscopy, GC-MS: gas chromatography-mass spectrometry, LC-MS: liquid chromatography-mass spectrometry. [Adapted from (Mussap et al., 2013)]

Technique	Benefits	Drawbacks
NMR	- Results in a fast and single experiment (*ca.* 10-30 min for 1D ^{1}H NMR) - Various nuclei (^{1}H, ^{13}C, ^{31}P) - Liquid and solid matrices - Very high accuracy and repeatability - Quantitative - Minimal sample preparation and sample recovery - Molecular dynamics information	- Needs very skilled technicians - Limitations in sensitivity (sub-mM) - NMR spectrometers are very expensive
GC-MS	- Very small amount of sample - High discrimination of molecules with a very similar structure - Very high sensitivity (< μM)	- Extensive sample preparation, including derivatization - Destructive - Limitations to volatile compounds
LC-MS	- Suitable for measuring lipids, di- and tripeptides and other macromolecules - Very high sensitivity (< μM)	- Extensive sample preparation - Destructive - Long analytical time (20-60 min per sample)

Both NMR and MS datasets of biological complex samples are usually vast, requiring the use of pre-processing and data mining methodologies based on multivariate analysis (MVA), for their handling and interpretation in terms of potential metabolite biomarkers (Rousseau et al., 2008). In parallel with MVA, the potential biomarkers are commonly inspected using univariate statistical analysis, followed by validation tools to check their accuracy. Biomarker identification is a very challenging step due the molecular complexity of biological samples. For NMR data, metabolite identification is achieved using databases of standard compounds and 2D NMR spectra. Metabolite identification in MS data is difficult due the lack of databases, and often extra experiments are needed such as tandem MS. At the end, the metabolite alterations are interpreted in terms of disturbed metabolic pathways (Lindon and Nicholson, 2008; Lämmerhofer and Weckwerth, 2013).

1.3.2. Brief state of the art of metabolomics

Regarding metabolomics studies, the most initial application was in drug toxicology, studying the effects of drugs on animal models by the measurement of alterations in the metabolic profile of biofluids (e.g., plasma and urine) over time, in relation to drug exposure (Nicholson et al., 2002). In this respect, metabolic nuances and biomarkers can be identified by characteristic changes in the pattern of concentrations of endogenous metabolites that relate to the site and mechanisms of toxicity.

The metabolomics strategy has also been applied in nutrition research focused on the study of the metabolic effects of nutrients or dietary patterns (e.g., chocolate intake) or nutrition-related diseases (e.g., obesity) on the metabolism (Rezzi et al., 2013). In addition, the number of publications of metabolomics applied to human disease research has been increasing in the last ten years, with the aim to derive new biochemically based assays and identify combination biomarkers for disease diagnosis, prognosis and management. For instance, there is a great need for sensitive and objective testing methods for cancer within the asymptomatic population, hence much research has been carried out in relation to colorectal, breast and lung cancers, among others (Duarte and Gil, 2012; Duarte et al., 2013). Cardiovascular diseases have also been investigated due to their high incidence of mortality in developed countries (Shah et al., 2012), as well as associated risk factors such as hypertension, type 2 diabetes and dyslipidemia (Rankin et al., 2014). Both NMR- and MS-based metabolomics have been used to comparable extents in disease research using mainly human samples. The study of prenatal disorders by NMR- and MS-based metabolomics has been

extensively addressed and a review of the main results reported so far is presented in the subchapter 1.4.

In relation to lipidomics, over the years, an increasing number of disorders have been linked with disturbances in lipid metabolism, thus lipidomics has been used to search for biomarkers, obtain knowledge on disease mechanisms and monitor the efficacy of therapeutic intervention (Vaz et al., 2015). Lipids have been broadly defined as hydrophobic or amphipathic small molecules that originate entirely or in part by carbanion-based condensations of thioesters and/or by carbocation-based condensations of isoprene units (Fahy et al., 2009). In blood plasma/serum, lipids are generally packed in particles called lipoproteins formed by an envelope containing apolipoproteins and amphipathic lipids (phospholipids and free cholesterol), and a core composed mainly by triglycerides, cholesteryl esters and small amounts of other lipids (Nelson and Cox, 2005).

In 2010, the human plasma lipidome was qualitatively and quantitatively described using several analytical techniques (GC-MS, LC-MS, MS/MS) unveiling six main categories (fatty acyls, glycerolipids, glycerophospholipids, sphingolipids, sterols and prenols), overall containing thousands of distinct lipid molecular species (Quehenberger et al., 2010). NMR is a less explored technique in lipidomics approach due to their broad and highly overlapped resonances, although some work has employed this technique to quantify lipoprotein particles in plasma, either using isolated fractions (by ultracentrifugation) or whole plasma/serum (Otvos et al., 1992; Jeyarajah et al., 2006; Ala-Korpela et al., 2007). The research in this area resulted in the creation of the *NMR LipoProfile® test* (LipoScience, Inc.), a novel NMR clinical application currently used in USA for measurement of lipoprotein particles in blood and individual lipids (triglycerides, cholesterol). In addition, few studies employed NMR spectroscopy to study the plasma/serum lipid profile using total lipid extracts (Willker and Leibfritz, 1998; Oostendorp et al., 2006; Tukiainen et al., 2008; Srivastava et al., 2010).

In clinical research, lipidomics has been extensively used in the study of obesity, diabetes, cardiovascular diseases, cancer, inflammations (e.g., rheumatoid arthritis), dysfunctions (e.g., mitochondrial dysfunction), genetic disorders (e.g., Barth syndrome) and Alzheimer's disease (Rolim et al., 2014). In prenatal research, a limited number of lipidomics studies have been published as will be further described in the subchapter 1.4.

1.3.3. Analytical techniques: Nuclear Magnetic Resonance (NMR) Spectroscopy and Mass Spectrometry (MS)

Principles of NMR spectroscopy

NMR is a spectroscopic technique based on the magnetic properties of some atomic nuclei when placed in a strong magnetic field, and provides extensive information on the molecular structure and dynamics of molecules (Hore, 1999). Atomic nuclei have a nuclear spin quantum number, I, which may be zero, a positive integer or a half integer. Only nuclei with a positive integer or half integer spin quantum number are "NMR visible", and these include ^{1}H, ^{13}C, ^{15}N, ^{19}F, ^{31}P ($I = 1/2$); ^{2}H, ^{14}N ($I = 1$); or ^{17}O ($I = 5/2$) (Claridge, 2009). The spinning nuclei possess angular momentum (P) and a magnetic momentum (μ) originating from their rotating movement, the spin. When exposed to an external static and homogeneous magnetic field, B_0, the magnetic moment of the nuclei experiences a torque which forces it into precession around the axis of the magnetic field (Figure 1.5a). In the case of ^{1}H (and any other ½ spin nucleus), there are two possible spin orientations or two different energy states. It can only align parallel (lower energy, α) or anti-parallel (higher energy, β) with respect to the external field, B_0, corresponding to magnetic quantum numbers of $m_1 = +1/2$ and $m_2 = -1/2$, respectively (Claridge, 2009). The torque imposed on the magnetic momentum due the presence of the static magnetic field (B_0) is known as Larmor precession, and is given by $\omega = -\gamma B_0$ (rad s^{-1}) or $\upsilon = -\gamma B_0/2\pi$ (Hz), where γ is the magnetogyric ratio of the nucleus (γ = 2.674×10^{8} rad·T^{-1} s^{-1} for ^{1}H). The difference of potential energy between the two spin states is given by $\Delta E = h\upsilon$.

Considering a collection of similar spin ½ nuclei in the applied magnetic field (Figure 1.5b), the parallel α orientation has slightly lower energy than the anti-parallel β orientation originating an excess of nuclei in the α state at equilibrium, as defined by the Boltzmann distribution (Equation 1.1), where N_α represents the number of nuclei in the parallel α spin orientation, N_β represents the number of nuclei in the anti-parallel β orientation, K_B the Boltzman constant and T the temperature.

$$\frac{N_\alpha}{N_\beta} = e^{\Delta E/K_B T} \qquad \text{Equation 1.1}$$

If the magnetic moments of all nuclei present in the sample are added vectorially, the difference of populations between α and β states for a nucleus with spin ½ leads to a net vector aligned with the +*z* direction, named bulk magnetization vector, M_0 (Figure 1.5b). In order to obtain a NMR signal, M_0 is flipped orthogonally to B_0 by use of a

radiofrequency (RF) field (90º pulse), B_1, oscillating at the same Larmor frequency as the spin, and applied for a defined time period. Consequently, M_0 is flipped from the z axis to the *x'-y'* plane and precesses with a frequency equal to the Larmor frequency ($\upsilon = -\gamma B_0/2\pi$ in Hz).

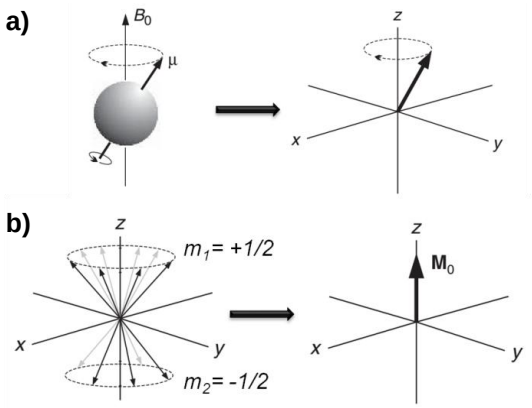


Figure 1.5. a) Precession of the magnetic moment $\boldsymbol{\mu}$ in the presence of an external magnetic field $\boldsymbol{B_0}$; b) at equilibrium the excess of spins in the $\boldsymbol{\alpha}$ state places the bulk magnetization vector $\boldsymbol{M_0}$ parallel to the +*z* axis. [Adapted from (Claridge, 2009)]

Upon energy absorption, there is an exponential decrease of the magnetization which induces a current in the receiver coil that also decays exponentially (Figure 1.6a). This process is recorded as a time domain $f(t)$ signal named Free Induction Decay (FID), which is then transformed into the frequency domain $f(w)$ by applying a Fourier transformation (Figure 1.6b) (Balci, 2005; Lindon et al., 2007).

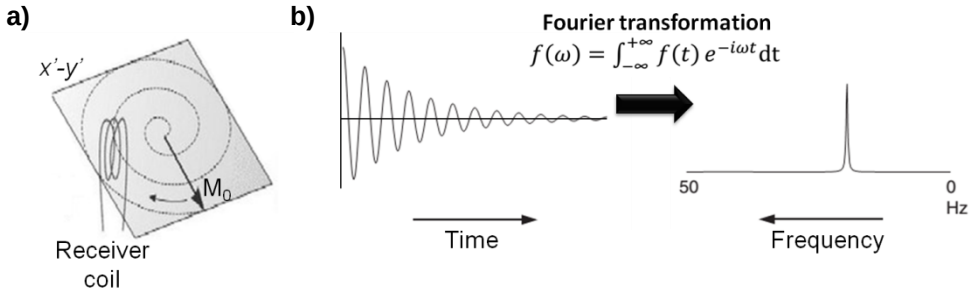


Figure 1.6. a) Scheme of the magnetization vector $\boldsymbol{M_0}$ precessing on the *x'-y'* plane around the z axis inducing a current in the receiver coil; b) Fourier transformation of time domain FIDs to the corresponding frequency domain spectrum. [Adapted from (Claridge, 2009)]

The behaviour of the spins during the period of free precession that follows the RF pulse is known as relaxation. There are two main types of relaxation processes, namely spin-lattice or longitudinal (T_1) and spin-spin or transverse (T_2) relaxation. The

T_1 is the time constant for the return to equilibrium of the populations of the spin states, while T_2 is the time constant for the dephasing of coherence between the spin states. Hence, the T_1 relaxation time characterizes the relaxation of longitudinal components of the magnetization M_0 parallel to B_0, or in the z component of the magnetization (M_z), described by Equation 1.2. The T_2 relaxation time characterizes the relaxation of transverse components of the magnetization M_0 perpendicular to B_0, or in the x-y component of magnetization ($M_{x,y}$), being given by Equation 1.3. T_2 is less than or equal to T_1, since return of magnetization of the z direction inherently causes loss of magnetization in the x-y plane. (Claridge, 2009)

$$M_z\,(t) - M_0 \propto e^{-t/T_1} \qquad \text{Equation 1.2}$$

$$M_{x,y}\,(t) \propto e^{-t/T_2} \qquad \text{Equation 1.3}$$

The transverse relaxation processes influence the widths of NMR resonances, the widths being inversely proportional to T_2 relaxation times. Molecules with a short T_2 are characterized by faster relaxation (fast decaying in FID) and broad resonances, whereas those of long T_2 originate slower relaxation (longer FIDs) and narrower resonances (Figure 1.7).

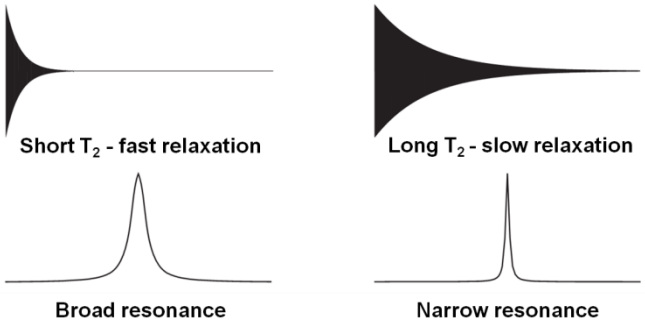


Figure 1.7. Influence of transverse relaxation (T_2) on the observed FID and NMR resonance width. [Adapted from (Claridge, 2009)]

Each nucleus is surrounded by electrons creating their own local magnetic field (B_{loc}) that opposes to the external magnetic field (B_0), thus creating a shielding effect in the nucleus. The effective field (B_{eff}) experienced by the nucleus is given by $B_{eff} = B_0 + B_{loc}$. Consequently, each nucleus resonates at a different frequency depending on the electronic/chemical neighbourhood, leading to different positions in the NMR spectrum. The chemical shift (δ) of a spin is always dimensionless and independent of B_0, given by the resonance frequency of a given nucleus (υ), the resonance frequency of a standard signal (υ_{stand}) and the frequency of the spectrometer (υ_{spec}), as described in Equation 1.4, expressed in parts per million (ppm) (Balci, 2005).

$$\delta = 10^6 \times \left(\frac{\upsilon - \upsilon_{stand}}{\upsilon_{spec}}\right) \qquad \text{Equation 1.4}$$

The most commonly used reference compounds for ^{1}H and ^{13}C are tetramethylsilane (TMS) and 3-(trimethylsilyl)propionate (TSP), which are set to zero. The range in which most of the ^{1}H NMR absorptions occur is relatively narrow, δ 0-14 downfield from the absorption of TMS. The general chemical shift ranges of different functional group can be seen in Table 1.7.

Table 1.7. Typical ^{1}H and ^{13}C chemical shifts of some chemical groups. [Adapted from (Fan, 1996)]

Chemical group	δ ^{1}H /ppm	Chemical group	δ ^{13}C /ppm
C-CH_3	0.7-1.1	-CH_3 (methyl)	8-30
=C-CH_3	1.5-1.8	-CH_2 (methylene)	14-55
$COCH_3$ (acetyl)	2.0-2.5	-CH- (methine)	22-60
N-CH_3	2.5-3.3	-C- (quaternary)	30-40
O-CH_3	3.3-4.3	CH_2=C-R	100-150
C-CH_2-C	1.3-2.5	C=C-C=C-R	110-150
CO-CH_2-C	2.5	C=C-R	65-90
O-CH_2-C	4.0-5.0	Heteroaromatic ring	100-165
C-CH-OH	3.3-4.0	C-OH (alcohol)	44-85
C-CH-O-ester	4.2-5.3	C-O-C (ether)	55-85
=CH-C (olefinic)	5.5-8.5	R-COOH (saturated)	165-188
Aromatic ring	6.0-9.0	R-COO^-	175-195
-CHO (aldehyde)	9.0-10.2	R-C=C-COOH	158-174
-COOH	10.5-13.5	R-COOR' (saturated)	158-178
C-OH (alcohol)	1.5-6.0	R-C=C-COOR'	152-172
C-OH (phenol)	6.5-18.5	R-CHO (saturated)	196-220
Primary amines	1.1-1.8	R-C=C-CHO	176-195
Secondary amines	1.2-2.1	Saturated ketones	
CO-NH-	5.0-6.5	R-CO-NH_2	150-178
N-alkyl amides	6.0-8.2		

The area of an NMR resonance is proportional to the relative number of nuclei giving rise to it, thus enabling quantitative measurements (Claridge, 2009). In order to get accurate NMR peak areas for quantification, the relaxation delay should be at least 5 × T_1 for the slowest relaxing ^{1}H in the sample. Typically, the concentration of a given compound is determined against that of a suitable reference compound (e.g., 3-(trimethylsilyl)propionate (TSP)) added to the sample. However, in some cases, the use of an internal chemical reference may impose constraints such as the interaction with the sample components (e.g., interactions between TSP and proteins in plasma/serum) which leads to overlap with compound resonances in the sample. An alternative approach is to introduce a synthetic rf reference signal as in the ERETIC

(Electronic Reference To access In Vivo Concentrations) method. In this method, the intensity of the synthetic signal is calibrated by a chemical reference solution of known concentration, defining the effective concentration of the ERETIC signal (Claridge, 2009).

One-dimensional (1D) and two-dimensional (2D) NMR spectroscopy in metabolomics

NMR spectroscopy has the extraordinary ability to provide detailed information on molecular structure of both pure compounds and complex mixtures such as foods or biofluids. However, the assignment of ^{1}H NMR spectra of a mixture is a complex procedure, and usually requires several 1D and 2D NMR methods, as well as, comparison with spectra of standard compounds. ^{1}H NMR experiments are largely used in metabolomics due the rapid acquisition times (few minutes per sample), making it suitable in studies with large sample arrays. Samples like biofluids have a high content of water and, hence, the water protons signal dominates the spectrum obscuring a large section. Thus, the water signal has to be suitably suppressed which is commonly achieved by water suppression pulse sequences such as a standard water presaturation, 1D noesy-presat, WATERGATE (water suppression by gradient-tailored excitation) or WET (water suppression enhanced through T_1 effects) sequences (Lindon et al., 2007). Usually, three types of 1D experiments are acquired in metabolomics studies of complex mixtures: (i) a 1D pulse-and-acquire experiment, usually referred as standard 1D, in which large and small molecules contribute to the spectrum with intensity proportional to their concentration; (ii) a transverse relaxation (T_2)-edited or Carr-Purcell-Meiboom-Gill (CPMG) experiment, in which signals from protons of macromolecules are suppressed resulting in a spectrum which is dominated by the smaller molecules; and (iii) a diffusion-edited experiment based on the difference in the diffusion coefficients between macromolecules and low molecular weight (M_w) metabolites, an approach in which only macromolecules (or bound small molecules) contribute to the spectrum (Beckonert et al., 2007; Lindon et al., 2007). The standard 1D pulse sequence consists of an RF 90º pulse preceded by a presaturation pulse to saturate the water signal. Usually, the standard 1D ^{1}H NMR spectrum of complex samples shows resonances of low M_w (e.g., glucose, lactate, amino acids) and high M_w metabolites (e.g., proteins, lipids, lipoproteins) heavily overlapped. The CPMG experiment uses the differences in T_2 relaxation times to attenuate the signals from macromolecules, and the resulting spectrum is dominated by small molecules (Meiboom and Gill, 1958; Claridge, 2009). Figure 1.8a shows a CPMG pulse sequence, where usually water attenuation occurs during relaxation delay, followed by

a 90º pulse, and a spin-echo sequence $[\tau\text{-}180°\text{-}\tau]_n$ used to selectively detect sharp signals from low M_w molecules. During the spin-echo delay (τ), the magnetization decreases exponentially with the ratio between τ/T_2, enabling the attenuation of macromolecules due their shorter relaxation times compared to small molecules (long T_2). In order to acquire the whole decay process a large number of echos (*n*) is needed. Diffusion-edited experiments are used to detect broad signals from macromolecules due their slow diffusion coefficients, attenuating sharp signals from small molecules (fast diffusion coefficients). Figure 1.8b shows a stimulated echo sequence for diffusion measurement with bipolar gradients, where an initial 90º pulse is applied followed by the application of gradients (G1) which diphase magnetization. During the diffusion delay (Δ), the molecules are allowed to diffuse, after which large molecules remain in a fairly constant position contrarily to small ones due the differences in the molecular size. The application of a subsequent gradient, opposite to the initial, refocuses magnetization of macromolecules and attenuate signals from low M_w molecules. Other pulse sequences commonly used for diffusion measurement are based in the pulsed field gradient spin-echo (Claridge, 2009).

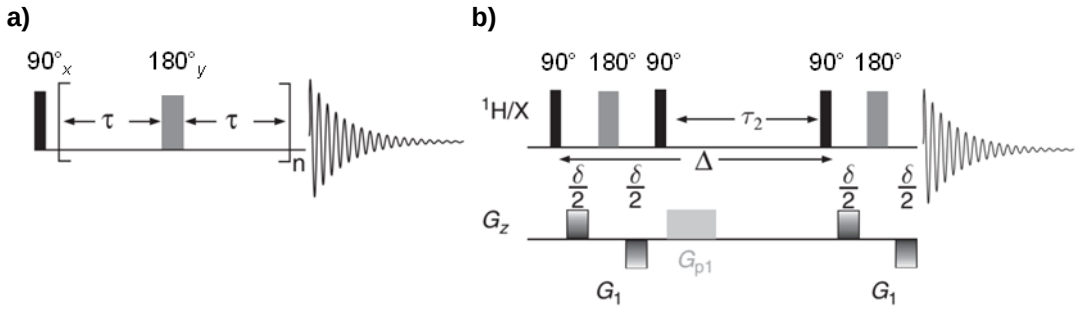


Figure 1.8. Pulse sequences of a) Carr-Purcell-Meiboom-Gill (CPMG) and b) diffusion-edited experiments. [Adapted from (Claridge, 2009)]

The use of 2D NMR methods is of paramount importance to provide additional information to solve overlap problems and allow identification of metabolites that otherwise would remain undetected. 2D NMR methods are based on dipolar (through space) or scalar (through bond) couplings between magnetic nuclei. 2D NMR acquisition has a general scheme with four basic steps (illustrated in Figure 1.9) corresponding to: 1) preparation period, spin systems relax and are excited by at least one RF pulse; 2) evolution period (t_1), when chemical shifts and spin-spin couplings evolve (time domain which is incremented in 2D experiments); 3) mixing period, where one or several RF pulses are applied and create an observable transverse magnetization; and 4) detection period (t_2), which corresponds to the acquisition time

(Claridge, 2009). Therefore, the acquisition of a 2D spectrum involves repeating a pulse sequence for increasing values of t_1 and recording a FID as a function of t_2 for each value of t_1 (Keeler, 2002b). The methods more commonly used for assignment and spectral simplification purposes in metabolomics are Total Correlation Spectroscopy (TOCSY), Heteronuclear Single Quantum Correlation (HSQC) and *J*-resolved spectroscopy (*J*-res), although two other experiments may also be helpful, namely Correlation Spectroscopy (COSY) and Heteronuclear Multiple Bond Correlation (HMBC) (Lindon et al., 2007).

The ^{1}H-^{1}H Correlation Spectroscopy (COSY) experiment reveals the network of spin-spin couplings (over 3-5 bonds) in each molecule and can provide simultaneously the chemical shift and scalar coupling by detecting the off-diagonal cross peaks, which are symmetric with respect to the diagonal (Lindon et al., 2007). The Total Correlation Spectroscopy (TOCSY) spectra are more information-rich than COSY since they detect multiple relayed connectivities, i. e., allow detection of long range ^{1}H-^{1}H spin-spin coupling connectivities (>3 bonds), that are usually too weak in COSY spectra (Lindon et al., 2007). TOCSY pulse sequence is characterized by an initial 90º pulse, followed by the evolution period (t_1) and the Malcolm Levitt (MLEV) sequence that allows coherence transfer between all coupled nuclei in a spin-system, as illustrated in Figure 1.9. The MLEV-17 sequence is formed by 16 composite pulses, followed by a regular uncompensated 180º pulse (Bax and Davis, 1985; Claridge, 2009).

In the Heteronuclear Single Quantum Correlation (HSQC) experiment, the chemical shift of ^{13}C is correlated via one-bond coupling constant $^{1}J_{CH}$ to the ^{1}H chemical shift of the directly bond proton (Lindon et al., 2007). The pulse sequence of HSQC, illustrated in Figure 1.9, shows a 90º pulse for proton magnetization followed by an insensitive nuclei enhancement by polarization transfer (INEPT) sequence, which transfers magnetization between the directly attached ^{1}H-^{13}C and also recovers magnetization from ^{13}C to ^{1}H before acquisition (Claridge, 2009). The Heteronuclear Multiple Bond Correlation (HMBC) experiment gives correlations between the ^{13}C nuclei and ^{1}H that are separated by two, three, and sometimes four bonds (Lindon et al., 2007).

The ^{1}H *J*-resolved spectroscopy (*J*-res) allows spectral simplification based on the spin-echo pulse sequence (Lindon et al., 2007). It separates *J*-splitting and chemical shifts on to two orthogonal axes: the *J*-multiplets are displayed along the F1 dimension, while the chemical shifts are shown in F2 dimension. This experiment minimizes overlap because contributions of substances with short T_2 values are suppressed, and allows better determination of coupling constants.

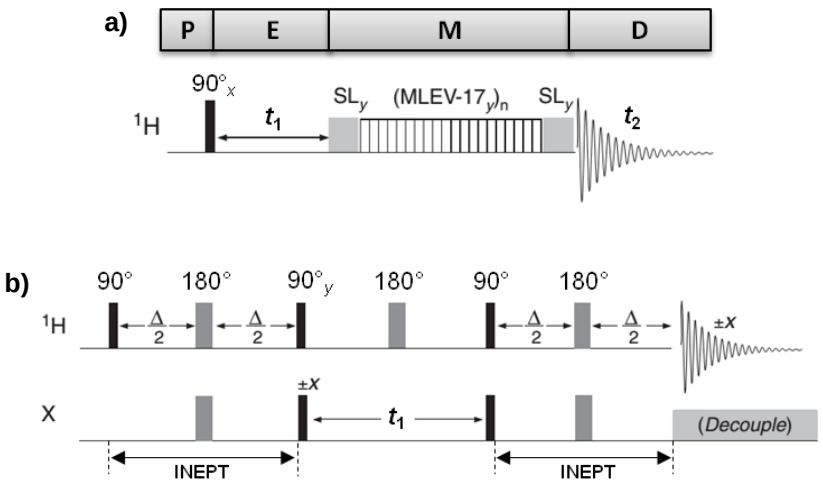


Figure 1.9. a) TOCSY pulse sequence based on the MLEV-17 mixing scheme. b) HSQC pulse sequence. P, preparation; E, evolution; M, mixing; D, detection. [Adapted from (Claridge, 2009)]

Other 2D experiment useful for assignment of complex mixtures is the diffusion ordered spectroscopy (DOSY). This experiment allows the discrimination of different components in a mixture by their translational diffusion coefficient as a consequence of their size (Kerssebaum, 2006). Some applications of this technique include determination of binding and association constants (Cohen et al., 2005), identification of unassigned spectral peaks in biofluids (Smith et al., 2007) and liquid foods (Gil et al., 2004), and determination of plasma lipoprotein sizes (Mallol et al., 2013).

The 2D NMR experiments described above are based on the excitation and detection of single-quantum coherence. A less explored 2D NMR technique in a metabolomics context but still explored for assignment of complex mixtures (e.g., poly- and monocyclic aromatic hydrocarbons, olive oil, polycyclic aromatic hydrocarbon pollutants) (Reddy and Caldarelli, 2010, 2011, 2012) is the Multiple-Quantum (MQ) spectroscopy. In MQ experiments, the excitation of MQ coherence is converted and detected as single-quantum coherence. The basic pulse sequence used to excite and detect MQ coherences is shown in Figure 1.10, where the first three pulses comprise the preparation period, and the last pulse is the mixing period. The first part of the sequence (90°-Δ/2-180°-Δ/2) is a spin echo which, at time (Δ), refocuses any evolution of offset frequencies but allows the coupling to evolve and generate anti-phase magnetization. The second 90° pulse turns the anti-phase magnetization into MQ coherence which, after evolving for time t_1, is returned into observable magnetization (single-quantum coherence) by the final 90° pulse (Keeler, 2002b).

The double quantum-single quantum (2Q-1Q) spectra are very similar to that of ^{1}H-^{1}H COSY enabling the identification of coupled spins without diagonal-peak multiplets. Series of MQ-1Q spectra can be acquired increasing the coherence order, thus reducing spectral complexity since some molecules are filtered out due to an insufficient size of their spin systems. As can be seen in Figure 1.10b, from a total of 11 compounds with similar molecular structures, only half appear in the 5Q-1Q spectrum. The signals observed in this spectrum correspond to the 5-proton molecular structures, thus it was possible to assign 6 compounds (phenol, *o*-terphenyl, dibenzyl, acetanilide, biphenyl and benzophenone) based on their fingerprint in 5Q-1Q experiment (Reddy and Caldarelli, 2010). In addition, the intensity of maximum quantum (MaxQ) signals can be quantified measuring the whole cross-peak 2D integral, thus enabling the quantification of molecules with very similar chemical structure (Reddy and Caldarelli, 2012).

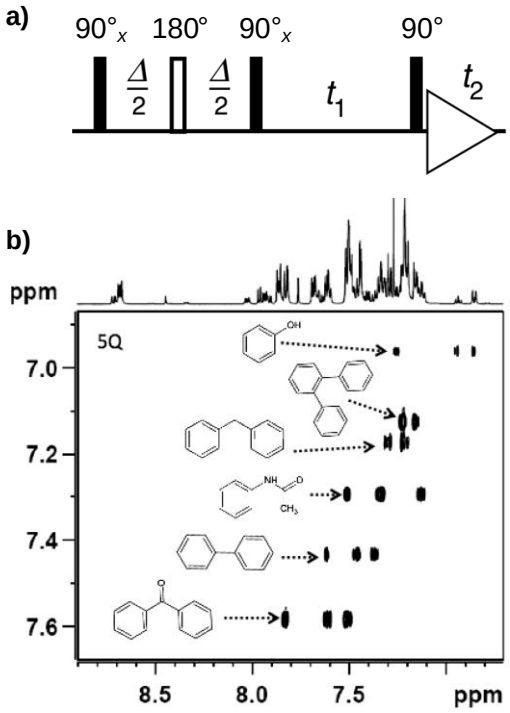


Figure 1.10. a) Pulse sequence for excitation and detection of MQ coherence. b) Expansion of the relevant aromatic region of the 5Q-1Q spectrum for assignment of phenol, *o*-terphenyl, dibenzyl, acetanilide, biphenyl and benzophenone in a mixture with 11 compounds. [Adapted from (Keeler, 2002b; Reddy and Caldarelli, 2010)]

Principles of MS

MS analysis consists in separation of gas-phase ions of compounds according to their mass-to-charge ratio (m/z). A mass spectrum (MS spectrum) plots *m/z* values of the ions against the relative abundance or ion intensity. Ions observed in the tandem mass spectrum (MS/MS spectrum) provide information about the nature and the

structure of their precursor molecule. So, the first step in MS is the production of gas-phase ions of the compounds under analysis in the ionization source. Under electrospray ionization, one of the most common ionization methods used nowadays, the molecule scan ionize as protonated molecules $[M+H]^+$. To gather structural information, these ions can be induced to fragmentation generating new ions, called product ions, and thus providing structural information about the initial molecule (Hoffmann and Stroobant, 2007).

A mass spectrometer essentially consists of five parts (Figure 1.11): 1) a sample inlet, to introduce the compound that is analysed (e.g. gas chromatograph, direct insertion probe), 2) an ionization source, to produce ions from the sample, 3) one or several mass analysers to separate the various ions according to their m/z, 4) a detector, to detect the ions emerging from the last analyser and measure their abundance, and 5) a data processing system, to produce the mass spectrum (Hoffmann and Stroobant, 2007).

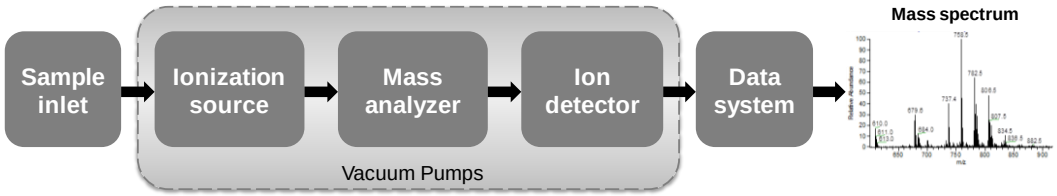


Figure 1.11. Main components of a mass spectrometer.

The sample inlet system can be direct (injects directly samples into the ion source), or alternatively coupled to a separation system such as GC or LC. General characteristics of GC and LC coupled to MS will be described below in the subsection *MS techniques in metabolomics and lipidomics*. A wide variety of ionization sources and mass analyzers are available, the choice being dependent on the nature of the sample and the type of information required from the analysis. The most common ionization sources are electron impact (EI), matrix-assisted laser desorption and ionization (MALDI) and electrospray ionization (ESI). Mass analysers include magnetic sector analyzers, quadrupole and ion trap analyzers (e.g. quadrupole ion trap (Q-IT)), time of flight analyzers (TOF), and Fourier transform ion cyclotron resonance. Following, the basic principles of the ESI and linear ion-trap are described, as they were the ionization source and analyzer used in this thesis. In ESI source, the sample solution is sprayed after passing through a needle and submitted to a high electric field, to create charged particles. Then, a flow of warm gas aid in solvent evaporation

allowing the formation of ions in the gas phase. Since ESI generates ions directly from solution, it is easily coupled with LC. Positive and negative single and/or multiply charge ions are created by ESI, depending on the nature of the analyte. Typically ESI-MS spectrum shows only the molecular ions with absence of fragment ions, since it is a soft ionization technique (Mishur and Rea, 2011). In the analyser linear ion trap, the molecular ions can be fragmented by collision with an inert gas, generating new ions that are further analysed and detected in a new MS/MS spectrum (Hoffmann and Stroobant, 2007). Linear ion trap is commonly used for analysis of low abundance compounds to their high trapping capacity as well as in fragmentation experiments to gain additional structural information (Lee et al., 2010). After sample ionization and sort by mass in mass analyzer, the detector converts the information to an electric current for quantitation and data analysis. A variety of detectors have been applied, the most widely used being the electron multiplier (Mishur and Rea, 2011).

In order to identify the molecular ions present in MS spectra, fragmentation experiments, known as tandem MS (MS/MS or MS^n), are widely applied enabling the comparison of the experimental fragmentation patterns with authentic standards and spectral data bases to confirm the molecular structure. In cases where data bases or authentic standards are not available, the metabolite identification is achieved using the exact mass of the precursor ion (unfragmented), the exact mass of the fragment ions, and their corresponding isotopic distributions (Lee et al., 2010).

MS techniques in metabolomics and lipidomics

Gas chromatography-mass spectrometry (GC-MS)

Metabolite profiling with GC-MS involves first an extraction step of metabolites from the biological sample, followed by derivatization to reduce the polarity and increase the thermal stability and volatility of metabolites making them amenable to GC. The derivatized molecules are injected into the GC system and subjected to physical separation in which the components are selectively distributed between the mobile phase (an inert carrier gas) and a stationary phase (thin film of liquid phase bonded to the walls). Capillary GC columns operate at very high temperatures and provide significantly high chromatography resolution. Chromatographic separation in GC is dependent of several factors such as column properties, carrier gas type, carrier gas velocity and oven temperature program (Pasikanti et al., 2008). Due the chemical derivatization required, GC provides separation of metabolites in a molecular weight range of 18 to ~350 Da and includes a range of relatively polar metabolite classes (e.g.

amino acids, organic acids, amines, amides, sugars, among others) (Dunn et al., 2011a). The ionization technique and mass analyzers mostly applied in GC-MS metabolite profiling are EI, and TOF or quadrupole analyzers, respectively. EI provides high reproducible fragmentation enabling the assignment of unknown metabolites using standard reference libraries (Want et al., 2006a; Dunn et al., 2011a). Recently, the introduction of comprehensive two-dimensional GC (GC×GC) combined with TOF-MS enable the reduction of analysis time as well as the detection of large number of metabolites in a single analytical run. GC×GC/TOFMS uses two capillary columns with different separation mechanisms (Dallüge et al., 2003; Pasikanti et al., 2008).

While in untargeted metabolomics the GC-MS apparatus is used for global metabolic profiling to cover all the measurable analytes in a sample (full scan mode), the strategy in targeted profiling is to select the ion signals specific to the chosen set of metabolites. The selection of specific ions is usually performed by selected ion monitoring (SIM). SIM is exclusively used to quantifying data on known targeted compounds providing increased sensitivity over full scan spectral acquisition used in untargeted profiling. In SIM mode, the mass spectrometer continuously measures only the selected molecule ions and characteristic fragments throughout the GC run (Hübschmann, 2009). Although LC-MS has been the most widely applied separation technique in lipidomics, some applications of GC-MS have been emerged. GC-MS has been mainly used to the analysis of fatty acids especially the cis/trans isomers (Kulig et al., 2006; Liu et al., 2010). The main limitations of GC for lipid analysis are the high temperatures used and the complexity of derivatization required before separation, being a time-consuming approach and removing most structural information about lipid molecular species.

Liquid chromatography-mass spectrometry (LC-MS)

When compared to GC, LC is not limited to the thermally stable volatile compounds and separation of polar and non-polar metabolites, salts, acids, bases and others is possible due the wide range of separation modes available. Reverse-phase LC (RPLC) is the most widely used LC method in untargeted metabolomics and lipidomics. However, hydrophilic interaction liquid chromatography (HILIC), particularly coupled to MS, has been emerged in these fields (Dunn et al., 2011a; Li et al., 2014). The RPLC method uses hydrophobic stationary phases (e.g., silica-based phases) and the mobile phases typically start with a high aqueous content (100% aqueous) to an increase of the organic phase fraction until the end (100% organic). This method provides efficient

retention and separation of relatively non-polar metabolites in a large M_w range (50 up to 1,500 Da) including high M_w lipid species (e.g., phospholipids and triglycerides) and non-polar amino acids among others (Dunn et al., 2011a). Compared to RPLC, HILIC enables the separation of more polar compounds typically non-retained in RPLC, thus combining both phases is possible to obtain a large coverage of the metabolites present in biofluids. HILIC stationary phases can be from three different types, namely neutral (no electrostatic interactions), charged (strong electrostatic interactions) and zwitterionic (weak electrostatic interactions), while the mobile phases usually are ammonium salts of acetic and formic acid at a suitable pH (Spagou et al., 2010). As an alternative to these techniques, ultra high performance chromatography (UPLC) has been applied. This technology combines high pressure with small particle size reverse phase packing material resulting in higher resolution, lower analysis time and increased sensitivity (Lindon et al., 2007).

In metabolomics and lipidomics, the ionization source most widely used with LC-MS systems is ESI, in both positive and negative modes, in order to obtain the most comprehensive profile. Regarding mass analyzers, quadrupole-TOF (Q-TOF) has been largely applied in LC-MS analysis of biofluids. This hybrid mass spectrometer combines high efficiency, sensitivity, and accuracy being typically coupled to ESI sources (Want et al., 2006a). In untargeted metabolic profiling, full scan acquisition is used in order to enable the detection of compounds of every mass and LC-MS/MS studies to provide structural information (Griffiths et al., 2010). Targeted studies using LC-MS also require appropriate LC separation conditions to the specific compound class of interest by performing selected reaction monitoring (SRM) scans determining precisely and accurately relative concentrations of the metabolites. In SRM, the fragmentation of the known precursor ion is monitored to a known fragmentation ion and several SRM experiments are undertaken during a single LC-ESI-MS/MS analysis to monitor a wide range of metabolites. LC-MS based lipidomics enables the analysis of almost all lipid molecular species with several studies performed on untargeted lipid profiling (Cífková et al., 2012; Sarafian et al., 2014), phospholipid profiling (Want et al., 2006b; Xia and Jemal, 2009; Scherer et al., 2010) and sphingolipid profiling (Hammad et al., 2010; Scherer et al., 2011) of blood samples.

Direct-infusion mass spectrometry (MS)

Alternatively to GC- and LC-MS, direct-infusion MS has been applied, without previous separation, to analyse all the lipid classes together. Lipid extracts are directly infused into the mass spectrometer where different ion source polarities (to form

positive or negative ions) and infused ionization solution additives, are employed to provide a lipid class-specific favoured analysis (Harkewicz and Dennis, 2011). ESI and matrix-assisted laser desorption/ionization (MALDI) are the most widely used ion sources in direct-infusion MS. Both ionization methods enable the measurement of many components simultaneously in complex lipid samples with good signal-to-noise ratios and reproducibility, without prior derivatization or separation. Targeted lipidomic analysis using ESI-MS specific scan models is globally referred as shotgun-lipidomics and has become the most frequently used MS approach in lipid profiling from complex samples such as blood and tissues (Li et al., 2014).

1.3.4. Statistical tools in metabolomics and lipidomics

Data pre-processing

The large and complex datasets generated by both NMR and MS-based techniques have to be first pre-processed for multivariate modelling. Usually, NMR data show variations in peak positions and intensity due to differences in sample matrix (i.e., pH or ionic strength, concentration). On the other hand, MS data commonly have background, noise and retention time fluctuations over a set of chromatograms. Therefore, pre-processing steps are very important to correct or minimize these variations. The pre-processing methods mainly used for both NMR and MS-based data are presented in Table 1.8.

Regarding the NMR data, the spectra can be divided into multiple discrete regions (buckets) along the chemical shift axis which are integrated, a process known as binning or bucketing. It was firstly applied to correct peak shifts due to pH or ionic strength variation among samples and reduce matrix complexity, however it frequently leads to substantial information loss (Blekherman et al., 2011). Alternatively, alignment algorithms can be applied to full resolution spectra in order to minimise the problem of peak shifts without decreasing spectral resolution. In the scope of this thesis, the recursive segment-wise peak alignment (RSPA) was applied (Veselkov et al., 2009). This algorithm consists of realigning peak positions in a ^{1}H NMR spectrum with respect to peak positions of a reference spectrum. It provides a fast and accurate spectral alignment using full spectral information and considering each individual signal as a whole and not as separate peaks. After either bucketing or alignment, a NMR two dimensional data table (or matrix) is constructed, consisting of *n* rows that correspond to observations/samples, and *m* columns that correspond to each variable (e.g., data point).

Table 1.8. Pre-processing methods mainly applied in both NMR and MS-based datasets.

NMR data	MS data
• Binning/bucketing	• Filtering
• Alignment	• Deconvolution
• Normalization	• Peak detection
• Scaling	• Alignment
	• Normalization
	• Scaling

Subsequently, to reduce the variation on overall concentrations of samples, a normalization step is usually performed. This step is crucial to remove unwanted systematic variation between samples. For NMR data, two types of normalization methods are commonly applied such as total area (or total intensity) or probabilistic quotient normalization (PQN). Total area normalization refers to divide each variable by the total integrated intensity across the whole NMR profile (Veselkov et al., 2011). PQN is based on the calculation of a most probable dilution factor by looking at the distribution of quotients of the amplitudes of a test spectrum by those of a reference spectrum (e.g., controls) (Dieterle et al., 2006).

The final pre-processing step is scaling which consists of compensating for the high dynamic range in metabolite concentration so that changes in abundant metabolites do not dominate statistical models. Scaling the data improves model interpretability, for example, unit variance (UV) consists in mean centering (i.e., subtracting the mean value of the sample) followed by the division of each variable by its standard deviation. Other scaling methods include pareto (division of each variable by the square root of the standard deviation of that variable) and logarithmic transformation (application of logarithm to each variable) (Veselkov et al., 2011). A more detailed description of pre-processing steps commonly applied in NMR data can be found in several review articles (Issaq et al., 2009; Blekherman et al., 2011; Liland, 2011).

For MS data, a higher number of pre-processing steps is required when compared to NMR. Spectral filtering is usually applied to remove random noise and baseline shift in the raw data (Katajamaa and Orešič, 2007). Then, spectrum deconvolution is used to separate two or more co-eluting or overlapping components in the MS spectral data. Peak detection enables the identification of all individual components caused by true ions and avoids detection of false positives providing accurate quantitative information. Chromatogram alignment consists of aligning the retention time to accurate *m/z*

ensuring that metabolites are identified as the same metabolite for all samples analysed, making the subsequent analysis more robust and accurate. The resulting dataset includes the retention time-m/z pairs, sample information, and peak areas.

The normalization of MS-based data is commonly achieved using either single or multiple stable internal standards with known concentration spiked into the sample prior to or during the extraction step (Blekherman et al., 2011). Finally, the scaling methods typically applied to MS data are the same enumerated above for NMR data. For those interested in a more detailed description of each MS data pre-processing step and corresponding algorithms, a number of reviews have been published (Katajamaa and Orešič, 2007; Issaq et al., 2009; Blekherman et al., 2011).

Multivariate analysis (MVA) tools: general considerations

Initially, unsupervised methods such as Principal Component Analysis (PCA) and hierarchical cluster analysis can be applied to NMR or MS-based datasets to reduce data complexity as well as visualize grouping trends and outliers (Lindon et al., 2001). PCA constructs a two-dimensional data table, called **X** matrix, containing *n* observations in a K-dimensional space (Figure 1.12a). The variation in the data table is reduced by a low dimensional model plane that approximates all rows (e.g. observations) in **X**, and is represented by points in Figure 1.12b. Hence, the variation in data matrix **X** is described by a scores matrix (T) which characterizes samples distribution, a loadings matrix (P) which gives the contribution of each variable to the samples distribution, and a residuals matrix (E) containing un-modeled variations, given by $\mathbf{X=TP^T+E}$. The 1st PCA component contains the largest matrix variance in the points. The 2nd PCA component is orthogonal to the first and contains the next largest variance and so on for a K-dimensional space. The scores ($\mathbf{t}_1$, $\mathbf{t}_2$, ..., $\mathbf{t}_m$) can be used to represent the relationship among the observations/samples in the model plane and provides an overview of all observations in the data table. Hence, groupings, trends, and outliers are made visible. The loading vectors ($\mathbf{p}_1$, $\mathbf{p}_2$, ..., $\mathbf{p}_m$) characterize the relationship among the measured variables and describe the contribution of those variables to the observed sample distribution. A central feature is that directions in the scores plot correspond to directions in the loadings plot, and vice-versa (Trygg et al., 2007).

Then, if needed, supervised methods such as partial least squares (PLS) and orthogonal PLS (OPLS) can be applied using the class information given for the training set of sample data to optimize the separation between two or more sample classes (Lindon et al., 2004). While PCA is a method that only considers the structure

of one matrix **X**, PLS constructs a set of orthogonal latent vectors from a matrix **X** (predictor variables) that maximizes the covariance between those vectors and a matrix ***Y*** (dependent variables), given by $\mathbf{X=TP^T+E}$ and $\mathbf{Y=TC^T+F}$.

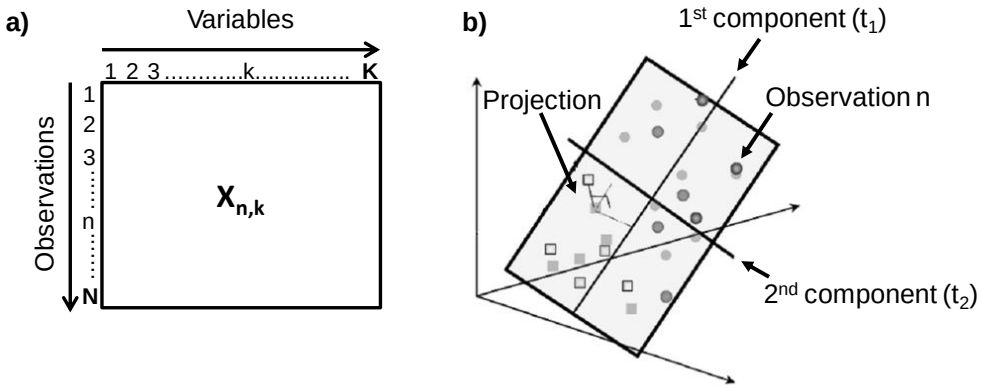


Figure 1.12. a) Data organization as a matrix denoted as X, with N rows of observations and M columns of variables; b) Geometrically, the principal components are equivalent to a plane, which well fits the N points in the K-dimensional space. [Adapted from (Trygg et al., 2007)]

PLS can be used for discrimination analysis, known as PLS-DA, where the **Y** matrix contains dummy values as class labels. In the case of two classes, usually the values of the dependent variable are given 1 for one class and 0 for the other class (Barker and Rayens, 2003). The metabolites contributing the most to the discrimination between the two classes are usually ascribed by plotting the loadings plots along with the variable importance to the projection (VIP). The metabolites with large VIP values in the model (higher than 1) are the most relevant and pointed as potential biomarkers (Mahadevan et al., 2008).

On the other hand, OPLS method is often applied with the aim to improve interpretation of PLS models and reduce model complexity. OPLS removes variation from **X** that is not correlated to **Y** or, in other words, removes systematic variation in **X** that is orthogonal to **Y** (Trygg et al., 2007). Other supervised methods that are often used are linear discriminant analysis, soft independent modelling of class analogy and neural networks (Lindon et al., 2001).

Variable selection

The data sets obtained from NMR and MS based metabolomics contain a large amount of information with some irrelevant or redundant variables for classification purpose. Therefore, the number of variables can be reduced using variable selection methods in order to improve model performance and reduce model complexity with

minor loss of information (Smolinska et al., 2012a). Variable selection methods have been applied to obtain a small set of variables in both NMR and MS data of biological samples (e.g., plasma/serum, cerebrospinal fluid, urine, tissue) related with non-alcoholic fatty liver disease (Quintás et al., 2009), multiple sclerosis (Smolinska et al., 2012b), colorectal cancer (Nishiumi et al., 2012), and different prenatal disorders (Diaz et al., 2013b). The variable selection techniques commonly applied in metabolomics are based on the multivariate analysis parameters, namely VIP values, loadings and PLS regression coefficients (b-coefficient) (Andersen and Bro, 2010; Smolinska et al., 2012a; Diaz et al., 2013b; Gromski et al., 2014).

One of the most used selection method has been based on the VIP criteria. The VIP reflects the importance of variables in the model both with respect to **Y** (i.e., its correlation to the response), and with respect to **X** (the projection). The VIP value for the k^{th} variable is given by Equation 1.5, where $\mathbf{W}_{kj}$ is the weight value for variable *k* component *j*, SSY_j is the sum of squares of explained variance for the j^{th} component, K is the number of variables, SSY_{total} is the total sum of squares explained of the dependent variable (**Y**), and J is the total number of components. A VIP value < 1 indicates a variable not relevant for explaining **Y**, which could be removed from the model (Andersen and Bro, 2010).

$$VIP_J = \sqrt{\frac{\sum_{f=1}^{F} W_{kj}^2 \cdot SSY_j \cdot K}{SSY_{total} \cdot J}} \qquad \text{Equation 1.5}$$

Other methods used in metabolomics studies are based on selectivity ratio, interval PLS (iPLS), genetic algorithms, Support Vector Machines-recursive feature elimination, random forests and recursive weighted PLS (rPLS). A detailed description of these methods can be found elsewhere (Rajalahti et al., 2009; Andersen and Bro, 2010; Lin et al., 2012; Gromski et al., 2014; Rinnan et al., 2014).

Validation of MVA models

When MVA is applied, the resulting models and respective discriminant metabolites have to be validated to ensure that the putative discovered biomarkers are not randomly correlated to the effect of interest. Usually, in order to determine whether overfitting occurred in MVA models, one of the most common methods used is *n*-fold cross-validation. Cross-validation consists of randomly spliting the dataset into *n* equally-sized blocks and then training the model *n* times each time holding out one of the folds as an internal validation set. The R^2 and Q^2 parameters are used to give

information about how well the data are modelled indicating the goodness of the fit and the predictive ability, respectively (Broadhurst and Kell, 2006). Thus, R^2 around 1 indicates perfect description of the data by the model, whereas Q^2 around 1 indicates perfect predictability. However, it has been shown that in practice it is difficult to give a general limit for a good Q^2 since this strongly depends on the properties of the dataset (e.g., number of observations included) (Triba et al., 2015).

Other common method to validate the MVA models is permutation testing. The permutation tests evaluate whether the specific classification of the individuals in the two designed groups is significantly better than any other random classification in two arbitrary groups. In this method, the two class labels are subjected to multiple random rearrangements of the labels on the observed data points. The permutation test is repeated many times and the permuted and true models are then compared to this distribution of all possible models (Westerhuis et al., 2008). In this thesis, Monte Carlo Cross Validation (MCCV) was computed for randomly permuted and true (original memberships) classes. MCCV consists in splitting the data into training and prediction sets, the latter set being used for class membership prediction. Usually, MCCV method performs 500 to 1000 iterations and calculates for each iteration the Q^2 value and confusion matrix, enabling at the end, the construction of a Q^2 distribution and a receiver operating characteristics (ROC) graph that give the prediction ability and the performance of the classification model (Figure 1.13) (Xu and Liang, 2001; Xia et al., 2013).

Confusion matrix represents four possible outcomes, namely true positives (TP), false negatives (FN), true negatives (TN), and false positives (FP), as shown in Figure 1.13a. It is the basis for several common metrics, such as classification rate (CR), sensitivity or true positive rate (TPR), and 1-specificity or false positive rate (FPR). The sensitivity (or TPR) is defined as the number of true positives as a percentage of all positives, given by Equation 1.6. 1-specificity (or FPR) is the number of false positives as a percentage of all negatives, defined by Equation 1.7. The classification rate (CR) is given by the ratio between the sum of true positives and negatives and the total number of classifications, as in Equation 1.8 (Fawcett, 2006).

The Q^2 distribution (Figure 1.13b) represents the Q^2 values obtained for permuted (H_0) and true models (H_A) in all iterations performed in MCCV, being a good manner to evaluate the predictive power of the model. Minimal overlapping of Q^2 values of true and permuted models usually indicates a model with good predictive power.

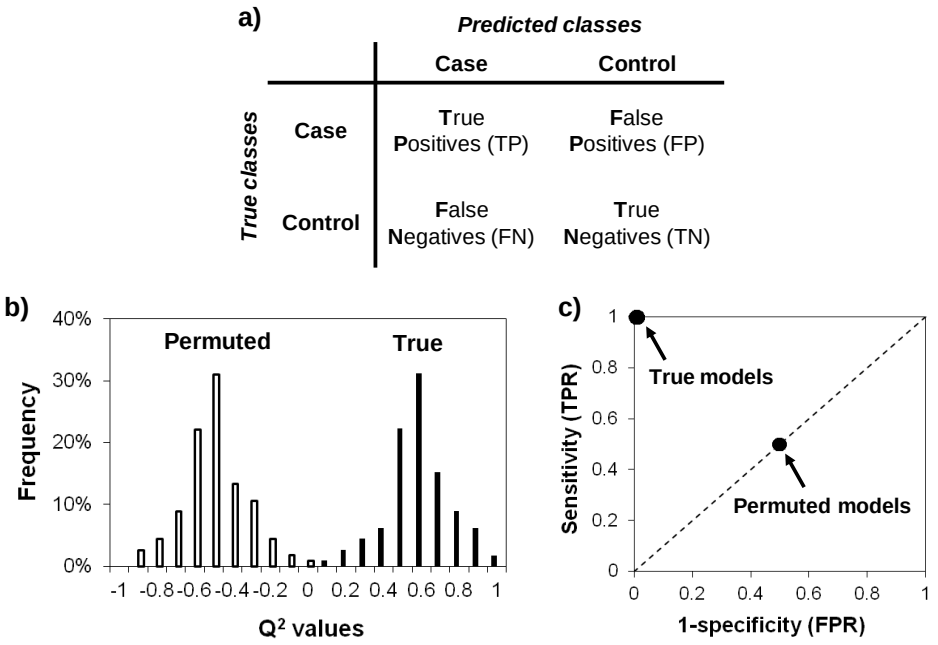


Figure 1.13. Examples of a) confusion matrix for two classes (disease and control), b) Q^2 distribution of true and permuted classes, and c) ROC graph of true and permuted models.

ROC graphs are constructed from the confusion matrix by plotting the sensitivity (TPR) on the *y* axis as a function of the 1-specificity (FPR) on the *x* axis. The ROC graph shown in Figure 1.13c shows the perfect classification since the true classes have sensitivity and specificity equal to 1 (100%) and are clearly separated from the permuted classes, which have both sensitivity and specificity equal to 50% (random classification performance). Alternatively, the area under the ROC curve (AUC) can be used to measure the accuracy of the classification model and usually ranges from 0 to 1. The AUC value goes to 1 for a perfect accuracy and separation between the classes. When two classes are not separated, the AUC value is close or lesser than 0.5 (Fawcett, 2006).

$$Sensitivity\ (TPR) = \frac{TP}{TP + FN} \qquad \text{Equation 1.6}$$

$$1 - specificity\ (FPR) = \frac{FP}{TN + FP} \qquad \text{Equation 1.7}$$

$$Classification\ rate\ (CR) = \frac{TP + TN}{TP + FP + TN + FN} \qquad \text{Equation 1.8}$$

Finally, the most reliable way to estimate the ability of the model to predict Y values of new individuals is to predict individuals from an independent sample set that were not used to build this model. If the same predictability appears in an independent study, the MVA models can be considered as reliable. However, in most metabolomics studies the number of samples is too low (dozens) to split the data into a training and a

test set, where the training set is used to build the model and the test set is used to estimate the predictability (Triba et al., 2015).

Statistical total correlation spectroscopy (STOCSY)

Statistical total correlation spectroscopy (STOCSY) has significantly enhanced information recovery from complex biological samples through the identification of multiple peaks from the same molecule, as well as the identification of different molecules involved in the same pathway (Cloarec et al., 2005). This method is based on the computation of correlation statistics between the intensities of all points in NMR spectra, thus generating links between signals from molecules that vary in concentration between samples (Cloarec et al., 2005; Coen et al., 2008). The application of STOCSY on metabolomic data was firstly reported by Cloarec et al. to decipher the structure of several metabolites in the ^{1}H NMR spectra of urine samples (Cloarec et al., 2005). This method has been applied in the metabolomics context for rapid identification of drug metabolites in biofluid samples (Holmes et al., 2007), integration of data from the same samples acquired on heteronuclear NMR experiments (e.g., ^{1}H and ^{31}P) (Coen et al., 2007) or different analytical platforms (e.g., NMR and ultra-performance liquid chromatography-mass spectrometry (UPLC-MS)) (Graça et al., 2012b), integration of data from the same individuals across different biofluids or tissues (Maher et al., 2011), and integration of data of genome-phenotype data with metabolomic data (Dumas et al., 2007).

The correlation matrix $\boldsymbol{C}$ is computed from a set of sample spectra according to Equation 1.9, where $\mathbf{X_1}$ and $\mathbf{X_2}$ are the autoscaled matrices composed by N spectra and K variables in the spectra of each matrix (Cloarec et al., 2005). The correlation coefficient gives information about the direction (positive or negative) and magnitude of the determined correlation ranging from 1 to -1.

$$\mathbf{C} = \frac{1}{\mathrm{n}-1}\mathbf{X_1}^{\mathrm{T}}\mathbf{X_2} \qquad \text{Equation 1.9}$$

STOCSY can be applied in both 1D and 2D forms and to homo (e.g., ^{1}H) or heteronuclear (e.g., ^{1}H and ^{31}P) data (Coen et al., 2008). The usefulness of STOCSY as a tool for molecular identification is clearly illustrated in Figure 1.14a, where all the ^{1}H NMR resonances from plasma glucose are highly correlated (red color). In addition, lower intermolecular correlations or anticorrelations represent molecules that may be involved in the same pathway, the concentrations of which are independent or under some common regulatory mechanism (Cloarec et al., 2005; Coen et al., 2008).

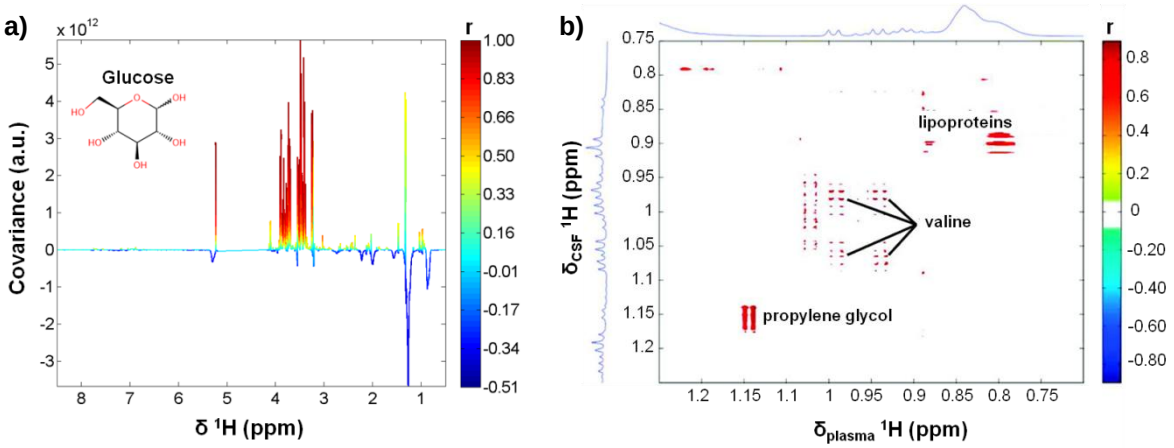


Figure 1.14. a) 1D STOCSY of the variable at δ 5.23 ppm correlated with the whole CPMG spectrum of blood plasma colored by the correlation coefficient (r); b) 2D STOCSY of CPMG 1H NMR spectra of blood plasma vs. cerebrospinal fluid (CSF). [Adapted from (Maher et al., 2011)]

The application of STOCSY to identify dynamic interactions between metabolites in two different biofluids, namely blood plasma and cerebrospinal fluid (CSF) (Maher et al., 2011), is illustrated in Figure 1.14b. This 2D plot shows high statistical correlations for valine between the two biofluids, along with propylene glycol, a drug delivery vehicle. In addition, correlations corresponding to broad signals from lipoproteins are observed.

Univariate statistical methods: general considerations

Data analysis methods in metabolomics are mostly based in multivariate analysis, although univariate methods have been used to identify biomarkers statistically significant according to a critical threshold (Vinaixa et al., 2012). These methods are based on a *null hypothesis* (H_0) which postulates a null difference between the mean (or median) of the areas or concentrations of the metabolites detected in the populations under study (e.g., controls and disease). The probability (p-value) for null hypothesis rejection is determined and if the p-value is below the pre-defined threshold of probability (α, usually set at 5%), the null hypothesis is rejected (Vinaixa et al., 2012). Univariate statistical methods can be parametric or non-parametric if the data follows a Gaussian or normal distribution or not, respectively (Vinaixa et al., 2012; Sheskin, 2014). Two types of parametric tests commonly used in metabolomics are Student's t-test and ANOVA (Analysis of Variance), while the Wilcoxon Rank Sum test is one of mostly used non-parametric tests.

1.4. Metabolomics and lipidomics in human prenatal and newborn health studies

The ability of metabolomics methods to detect simultaneous changes in the metabolites present in biological samples has advanced enticing possibilities for improved prenatal health monitoring. Figure 1.15 illustrates the number of papers published in relation to metabolomics (based on NMR, MS and infrared (MIR and NIR)) and lipidomics of prenatal disorders comprising maternal, fetal and newborn conditions. Most metabolomic studies have taken place in the last four years (Figure 1.15a) with the number of papers starting to increase from 2011 (coinciding with the beginning of this thesis) until 2014. Lipidomic studies comprise only 8 papers focused on MS techniques (GC- and LC-MS) with only one based in NMR applied to measure lipoprotein concentrations.

Regarding the biological matrices used in these studies, maternal blood (plasma and serum) has been the most explored biofluid, followed by amniotic fluid, maternal urine, umbilical cord blood, newborn blood and urine, placental extracts, breast milk, cervicovaginal fluid, bronchoalveolar lavage fluid, and hair, in decreasing order of number of publications (Figure 1.15b). For maternal blood plasma studies, three resulted from this thesis (Diaz and Pinto et al., 2011; Pinto et al., 2014b, 2014d). The most studied disorder was found to be preeclampsia, followed by PTD, FM, perturbations of fetal growth (including SGA, LBW, VLBW, IUGR and macrosomia), GDM, CD, healthy pregnancy, PROM, IEM and HIE, in decreasing order (Figure 1.15c). The increasing awareness of the scientific community towards the usefulness of metabolomics in prenatal research has also been illustrated by a large number of reviews published from 2006 to 2014 (one of them resulted from the work developed during this thesis (Pinto et al., 2014a)) listed in Table 1.9 by research area. A brief review of the main results reported for prenatal and newborn health studies is presented below, firstly for healthy pregnancies and then for maternal and fetal disorders. More recent studies comprising higher number of samples will be presented in more detail, grouped first by biological matrices near the fetal environment (amniotic fluid and placenta), followed by maternal (blood and urine) and newborn biofluids (umbilical cord blood, dried blood spots and urine).

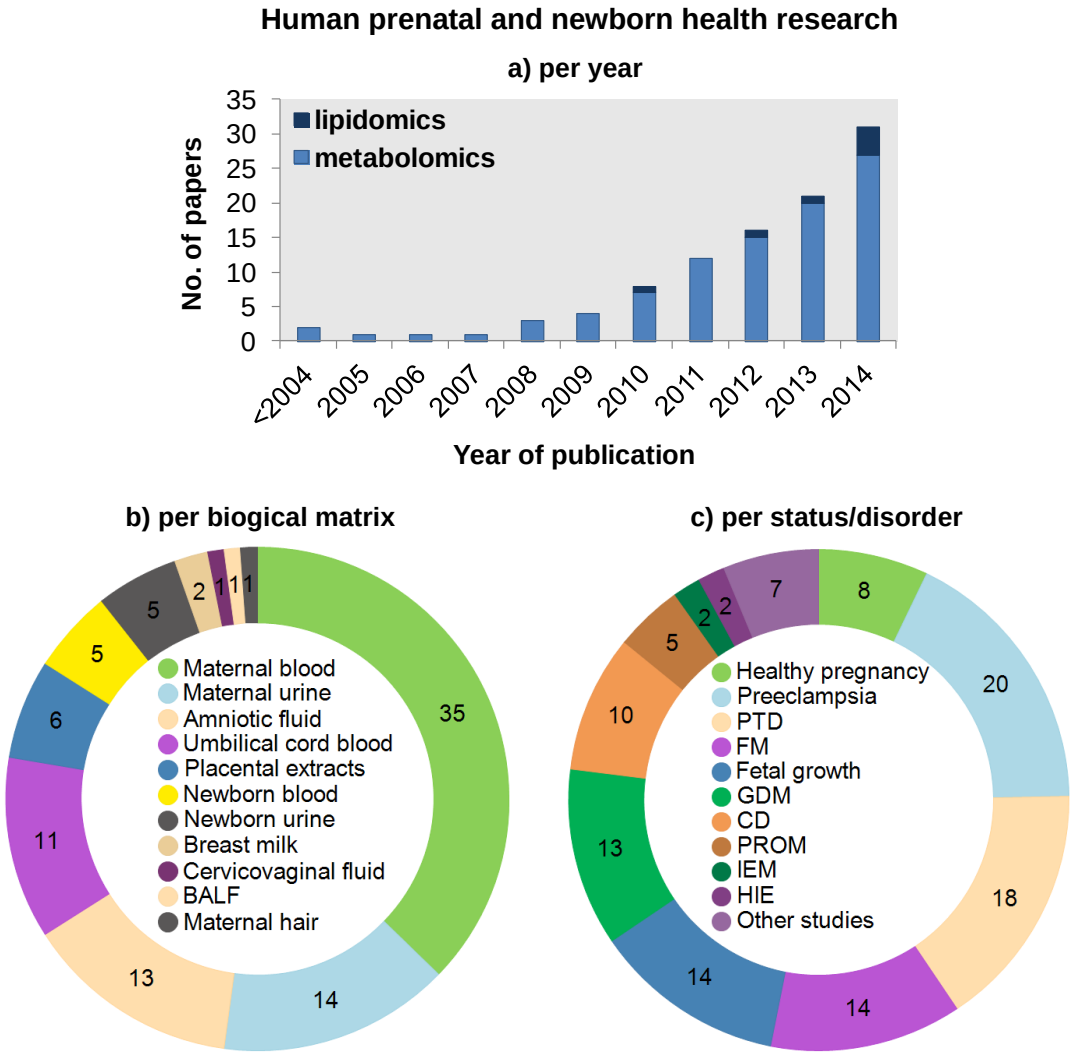


Figure 1.15. Number of metabolomics and lipidomics studies in human prenatal and newborn health research shown by a) year published (including reviews), b) biological matrix used and c) status (healthy or non healthy). BALF: bronchoalveolar lavage fluid, IEM: inborn errors of metabolism, HIE: hypoxic ischemic encephalopathy. Fetal growth anomalies include SGA, low birth weight (LBW), very low birth weight (VLBW), IUGR and macrosomia. Other studies include delivery method, newborn's age, biliary atresia and hepatitis syndrome, pesticide exposure, breastfeeding, cytomegalovirus, respiratory distress syndrome, and bronchopulmonary dysplasia.

Table 1.9. Metabolomics reviews of maternal, prenatal and newborn disorders.

Research area	Reference
Maternal disorders	(Romero et al., 2006), (Horgan et al., 2009), (Gracie et al., 2011), (Heazell et al., 2011), (Syggelou et al., 2012), (Fanos et al., 2013a), (Lowe and Karban, 2014), (Dessì et al., 2014a), (Huynh et al., 2014), (Pinto et al., 2014a)
Fetal disorders	(Dessì et al., 2012), (Dessì et al., 2013), (Fanos et al., 2013a), (Dessì et al., 2014a), (Kamath-Rayne et al., 2014), (Pinto et al., 2014a)
Newborn disorders	(Syggelou et al., 2012), (Fanos et al., 2013b), (Mussap et al., 2013), (Moco et al., 2013), (Dessì et al., 2014b), (Pinto et al., 2014a)

1.4.1. Healthy pregnancy and maternal disorders

Table 1.10 lists the metabolomics studies carried out on different biological matrices in search for metabolite alterations in relation to healthy pregnancy and maternal disorders. Healthy pregnancies have been studied by metabolomics of amniotic fluid (Cohn et al., 2009, 2010; Ottolenghi et al., 2010), placental tissue extracts (Dunn et al., 2011b), maternal blood plasma (Luan et al., 2014; Pinto et al., 2014b) and urine (Sachse et al., 2012; Diaz et al., 2013a; Trivedi and Iles, 2014a), one comprising results presented in this thesis (Pinto et al., 2014b). NMR and GC-MS metabolomics have revealed gestational age-related alterations in amniotic fluid levels of several amino acids (e.g., Ala, Val, Ile, Leu, Lys, Gln), betaine, creatinine, glucose, pyruvate and succinic acid (Cohn et al., 2009, 2010; Ottolenghi et al., 2010). For placental tissue extracts collected at early and late 1st trimester, significant changes in acyl carnitines, di- and triglycerides, nucleotides, fatty acids, phospholipids, sphingolipids and vitamin D metabolites were observed by UPLC-MS (Dunn et al., 2011b).

For maternal blood plasma, untargeted LC-MS and shotgun lipidomics of plasma extracts of 1st (n=30), 2nd (n=120) and 3rd (n=30) trimesters revealed increases in dihydrobiopterin, free carnitine, acetyl-alanine, leucyl-phenylalanine and in several phospholipid classes (phosphatidylcholines (PC), lyso-PC (LPC), lysophosphatidylethanolamines (LPE) and lysophosphatidylserines (LPS)) (Luan et al., 2014). Regarding maternal urine alterations across pregnancy, a large (n=823 women) and multi ethnic study was performed by NMR metabolomics including samples collected during pregnancy (8-20 g.w. and 26-30 g.w.) and post partum (10-13 weeks after delivery) unveiling substantial differences between the time points, dominated by a steady increase of urinary lactose concentrations, and an increase during pregnancy and subsequent dramatic reduction of several unidentified NMR signals between 0.5 and 1.1 ppm (Sachse et al., 2012). Other NMR metabolomics study of urine collected for non-pregnant (n=16) and pregnant at 1st (n=16), 2nd (n=20) and 3rd (n=19) trimesters revealed significant changes in 21 metabolites, 8 of which newly detected in relation to pregnancy: choline, creatinine, 4-deoxyerythronic acid, 4-deoxythreonic acid, furoylglycine, guanidoacetate, 3-hydroxybutyrate and lactate (Diaz et al., 2013a).

Table 1.10. Metabolomics studies of healthy pregnancy and maternal disorders by biological matrix. † lipidomics studies. NB: newborn; CVF: cervicovaginal fluid.

Matrix	Subject	References
Amniotic fluid	Healthy	(Cohn et al., 2009), (Cohn et al., 2010), (Ottolenghi et al., 2010)
	Preeclampsia	(Bock, 1994)
	PTD	(Graça et al., 2010), (Romero et al., 2010), (Power et al., 2011), (Graça et al., 2012a), (Graça et al., 2012b), (Graça et al., 2013), (Menon et al., 2014)
	GDM	(Bock, 1994), (Graça et al., 2010), (Graça et al., 2012a), (Graça et al., 2012b)
	PROM	(Graça et al., 2010), (Graça et al., 2012a), (Graça et al., 2013)
Placenta	Healthy	(Dunn et al., 2011b)
	Preeclampsia	(Heazell et al., 2008), (Dunn et al., 2009), (Dunn et al., 2011b), (Baig et al., 2013)†, (Korkes et al., 2014)†
Maternal blood	Healthy	(Luan et al., 2014)†, (Pinto et al., 2014b)
	Preeclampsia	(Kenny et al., 2005), (Turner et al., 2007), (Kenny et al., 2008), (Turner et al., 2008), (Kenny et al., 2010), (Odibo et al., 2011), (De Oliveira et al., 2012)†, (Bahado-Singh et al., 2012), (Bahado-Singh et al., 2013a), (Austdal et al., 2014), (Kuc et al., 2014), (Pinto et al., 2014d)†, (Korkes et al., 2014)†
	PTD	(Thorp et al., 2013)†, (Pinto et al., 2014d)†
	GDM	(Chen et al., 2010)†, (Diaz and Pinto et al., 2011), (Graça et al., 2012a), (Scholtens et al., 2014), (de Seymour et al., 2014)
	PROM	(Diaz and Pinto et al., 2011), (Graça et al., 2012a)
Maternal urine	Healthy	(Sachse et al., 2012), (Diaz et al., 2013a), (Trivedi and Iles, 2014a)
	Preeclampsia	(Diaz et al., 2013b), (Austdal et al., 2014)
	PTD	(Diaz and Pinto et al., 2011), (Graça et al., 2012a), (Diaz et al., 2013b)
	GDM	(Diaz and Pinto et al., 2011), (Graça et al., 2012a), (Graça et al., 2012b), (Sachse et al., 2012), (Diaz et al., 2013b), (Dudzik et al., 2014)
	PROM	(Diaz and Pinto et al., 2011), (Graça et al., 2012a), (Diaz et al., 2013b)
	Others	(Bonvallot et al., 2013), (Sachse et al., 2014)
Umbilical cord blood	GDM	(Dani et al., 2014)
	Delivery	(Hashimoto et al., 2013)
NB blood	PTD	(Wilson et al., 2014)
NB urine	PTD	(Atzori et al., 2011), (Moltu et al., 2014)
Breast milk	PTD	(Marincola et al., 2012), (Longini et al., 2014)
CVF	PTD	(Auray-Blais et al., 2011)

Preeclampsia

Preeclampsia has been the subject of most of the metabolomics work carried out so far, with studies performed on amniotic fluid by NMR (Bock, 1994) and placenta by GC- and UPLC-MS (Heazell et al., 2008; Dunn et al., 2009, 2011b) comprising samples collected after diagnosis (at delivery) in small cohorts (n ≤ 11). Metabolic

profiling of amniotic fluid by NMR revealed higher acetate, succinate, choline and citrate in preeclampsia cases compared to controls (healthy pregnant women) (Bock, 1994). Metabolomics studies (GC-MS and LC-MS) performed using placental extracts showed several alterations related with preeclampsia, suggesting oxidative stress and hypoxia (Heazell et al., 2008; Dunn et al., 2009, 2011b). In addition, lipidomics studies of placenta lipid extracts (preeclampsia cases n≤10, controls n≤10) unveiled alterations in total phosphatidylserines (PS), phosphatidylinositols (PI), phosphatidic acids (PA), and ganglioside mannoside 3 (Baig et al., 2013), as well as alterations in phosphatidylethanolamines (PE), phosphatidylglycerols (PG), diacylglycerols (DG), triacylglycerols (TG), acidic glycosphingolipids, and steroid conjugates (Korkes et al., 2014).

The first metabolomic studies of preeclampsia including maternal blood samples comprised small sample cohorts (n≤20) collected after diagnosis (> 20 g.w.) (Kenny et al., 2005, 2008; Turner et al., 2007, 2008). Thereafter the need for larger cohorts and sample collection before the onset of the disease (in 1st and early 2nd trimesters) was recognised in order to find potential predictive biomarkers. Kenny et al. found a discriminatory metabolite signature of pre-diagnostic preeclampsia (n=60 in test set and n=39 in validation set) composed by 14 metabolites (comprising amino acids, carbohydrates, carnitines, dicarboxylic acids, fatty acids, ketones, keto- and hydroxyl-fatty acids, lipids, phospholipids and steroids) in maternal plasma collected at 15±1 g.w. with an AUC of 0.92 by UPLC-MS (Kenny et al., 2010). Moreover, several attempts have been made to predict specific subphenotypes of preeclampsia, namely late- and early-onset preeclampsia which require delivery before 34 g.w and ≥34 gw., respectively (Bahado-Singh et al., 2012, 2013a; Kuc et al., 2014). Bahado-Singh et al. reported a combination of maternal serum metabolites (citrate, glycerol, hydroxyisovalerate and Met measured by NMR), with uterine artery pulsatility index and fetal crown-rump length which showed better sensitivity (82.6%) and false positive rate (1.6%) for early-onset preeclampsia (n=30) than the set of metabolites alone (sensitivity 75.9%, FPR 4.9%) (Bahado-Singh et al., 2013a). Other authors reported lower detection rate and FPR (55% and 10%, respectively) for early-onset preeclampsia (n=68) combining metabolites (taurine and Asn measured by UPLC-MS) with mean arterial pressure (Kuc et al., 2014). To date, none (nor any combination) has emerged with the requisite specificity and sensitivity to be of clinical use. Regarding the application of lipidomics approach to find lipid changes related with the onset of preeclampsia, the first one resulted from the work developed in this thesis with focus on plasma phospholipid profiling by HILIC-LC/MS (Pinto et al., 2014d). Later on, Korkes et al. found increases in PC, PS, PA, sphingomyelins (SM), glycosphingolipids,

and steroid conjugates, and decreases in PE, PG and sterols, in plasma (32-39 g.w.) of women with diagnosed preeclampsia (n=10) compared to controls (n=10), by MALDI-MS (Korkes et al., 2014).

The search for predictive biomarkers of preeclampsia in maternal urine is less explored comprising only two publications by NMR (Diaz et al., 2013b; Austdal et al., 2014). For pre-diagnostic preeclampsia apparently specific decreases in acetate, formate (Krebs cycle precursors) and fumarate (Krebs intermediate, along with 2-ketoglutarate and succinate, also decreased), along with a decrease in Ile were detected by NMR in a small cohort of women who later developed preeclampsia (n=9, collected at 16-17 g.w.) (Diaz et al., 2013b). Urine samples collected after the diagnosis of preeclampsia (22-38 g.w.n n=10) revealed, by NMR, increased choline and decreased Gly, p-cresol sulfate and hippurate, which may be related to increased oxidative stress and kidney dysfunction (Austdal et al., 2014).

Preterm delivery (PTD)

The metabolomics studies of PTD have been focused in the search for potential predictive biomarkers of the disease, collecting samples with absence of clinical symptoms. These studies have been performed using amniotic fluid revealing alterations in allantoin, citrate, amino acids, organic acids, carbohydrates and proteins, which suggest oxidative stress, early perturbations in placental-amino acid transport, hyperglycaemia and changes in protein metabolism (Graça et al., 2010, 2012b, 2013; Romero et al., 2010; Power et al., 2011; Menon et al., 2014).

The use of maternal plasma to study PTD is still very limited having been only explored by lipidomics for measurement of lipoprotein concentrations (Thorp et al., 2013) as well as the study of phospholipid profile (Pinto et al., 2014d) before the onset of the disease, the latter comprising results presented in this thesis (Chapter 8, section 8.3). Maternal plasma collected at 16-21 g.w. for 132 controls and 128 cases of recurrent PTD unveiled only alterations in VLDL particle size, with no changes in the remaining 16 lipoprotein classes measured by NMR spectroscopy (NMR Lipoprofile autoanalyzer) (Thorp et al., 2013). For maternal urine, the number of publications is also very limited, comprising one study by NMR which reported increased ketone bodies and decreased carnitine for pre-diagnostic PTD (n=26, collected at 15-25 g.w.), suggesting enhanced lipid oxidation and oxidative stress (Diaz et al., 2013b). Other studies aimed at analyse cervicovaginal fluid in women giving birth prematurely compared to term, by UPLC-MS with no further metabolite identification (Auray-Blais et al., 2011), and the metabolic profile of preterm breast milk by NMR and GC-MS

(Marincola et al., 2012). Urinary metabolic profiles of infants born preterm compared to term have revealed alterations in hippurate, Trp, Phe, malate, Tyr, hydroxybutyrate, N-acetyl-glutamate and Pro (Atzori et al., 2011).

Gestational Diabetes Mellitus (GDM)

The first study on GDM by a metabolomics-based approach was published for amniotic fluid using NMR (Bock, 1994). The most complete study was performed by Graça et al. using NMR and comprising samples collected before the onset of GDM (14-25 g.w., n=27) (Graça et al., 2010). Compared with healthy women (n=82), pre-diagnostic GDM women exhibited an average increase in glucose and small decreases in several amino acids along with acetate, formate, creatinine, and glycerolphosphocholine, suggesting hyperglycaemia, disturbances in amino acid and lipid metabolism, nucleotide biosynthesis and renal function. Later on, Graça et al. reported a new study to investigate the possible effects of pre-diagnostic GDM on the metabolic profiling of amniotic fluid viewed by UPLC-MS, revealing no significant alterations (Graça et al., 2012b).

Metabolomics was also applied to study the metabolite profile of maternal blood and urine in cases of pre-diagnostic GDM (Diaz and Pinto et al., 2011, 2013b; Graça et al., 2012a; de Seymour et al., 2014), post-diagnostic GDM (Chen et al., 2010; Sachse et al., 2012; Dudzik et al., 2014), and pregnant women with hyperglycaemia (Chen et al., 2010; Scholtens et al., 2014). The search for metabolic alterations in maternal plasma preceding the onset of GDM was first performed in a small cohort of pre-diagnostic GDM (n=19) women compared to controls (n=25) by NMR, in their 2nd trimester (16-22 g.w.) (Diaz and Pinto et al., 2011). Metabolite alterations comprise decreases in TMAO and betaine reflecting choline-related homocysteine-methionine conversion and possible alterations in renal function. Seymour et al. investigated the serum metabolite profile of pre-diagnostic GDM cases (at 20 g.w. for n=22) by GC-MS whose results unveiled only a significant increase for itaconic acid (de Seymour et al., 2014). Regarding post-diagnostic GDM, maternal plasma of pregnant women with GDM, collected at the day of OGTT (22-28 g.w., n=20), was compared with a control group (n=20) by a multiplatform system (LC-MS and GC-MS) (Dudzik et al., 2014). Results enabled the identification of some metabolites with a strong discriminative power for GDM (several lysophospholipids, taurine-bile acids and long-chain polyunsaturated fatty acid derivatives) indicative of low-grade inflammation and altered redox-balance. In addition, a targeted GC-MS study to serum fatty acids (FA) of controls (n=98), women with hyperglycaemia less severe than GDM (n=80), and

women with GDM (n=49), unveiled a graded increase in total FA and several individuals FA (e.g., linoleic, arachidonic, docohexaenoic) in the 3rd trimester (30 g.w.) (Chen et al., 2010).

Maternal urine collected for pre-diagnostic GDM (n=42) cases and controls (n=84) at 16-22 g.w. revealed several metabolite alterations determined by NMR namely a marked increase in glucose (possible early sign of glycemia), and other changes in common with type 2 diabetes (increased acetate, N-methylnicotinamide and N-methyl-2-pyridone-5-carboxamide, changes in creatine and creatinine and decreases in hippurate and phenylacetylglutamine) (Diaz et al., 2013b). The study of the urinary NMR profile of pregnant women diagnosed with GDM (three time points collection: 8-2 and 26-30 g.w. and 10-16 weeks post partum, n=823 women with different ethnicity) unveiled no reliability to identity GDM cases suggesting only an increase in median citrate concentration with increasing hyperglycaemia (Sachse et al., 2012).

An additional study performed the comparison of the NMR cord serum metabolite profile of infants of GDM mothers (IGDMs, n=30) to that of infants of healthy mothers (n=40), to evaluate if differences remain despite a strict control of GDM. Results showed lower glucose and higher levels of pyruvate, His, Ala, Val, Met, Arg, Lys, hypoxanthine, lipoproteins and lipids than controls (Dani et al., 2014). A recent review of metabolic profiling of GDM can be found elsewhere (Huynh et al., 2014).

Premature rupture of membranes (PROM)

The search for potentially predictive markers of PROM has also been attempted by analyzing amniotic fluid (Graça et al., 2010, 2012a, 2013), maternal blood plasma (comprising results presented in this thesis) (Diaz and Pinto et al., 2011) and maternal urine (Diaz and Pinto et al., 2011, 2013b) in the 2nd T (i.e. pre-diagnosis state). The NMR profile of amniotic fluid of 34 pre-PROM compared to 82 controls unveiled small metabolite changes (Met, Gln, Thr and two unassigned resonances at 3.11 and 3.29 ppm) (Graça et al., 2010). In a later study, Graça et al. explored the potential of MIR metabolic profiling of amniotic fluid to unveil differences between 37 pre-PROM and 40 controls, noting minor changes along with a low predictive power (Graça et al., 2013). The NMR metabolic profile of maternal urine was also explored to find differences related with pre-PROM (n=68) compared with controls (n=84) but without success since non predictive models were found (Diaz et al., 2013b). The low impact of PROM in amniotic fluid and maternal urine can be a) a direct reflection of the low gravity of PROM (after the 37th g.w.), as it is usually resolved by inducing labour without posing

significant risks for the mother and fetus, and/or b) the high interval from sample collection to diagnosis.

1.4.2. Fetal and newborn disorders

Table 1.11 lists the existing metabolomic reports on different biological matrices in relation to fetal and newborn disorders.

Fetal Malformations (FM)

Fetal disorders have an impact on fetal/maternal metabolism and the effects of fetal malformations (FM) on amniotic fluid metabolite profile have been investigated by NMR, UPLC-MS and MIR (Groenen et al., 2004; Graça et al., 2009, 2010, 2012a, 2012b, 2013). Groenen et al. evaluated the effect of spina bifida (n=14 at 15 and 39 g.w.) on amniotic fluid composition by NMR having found changes in succinic acid and Gln (increased levels) suggestive of derangement in amino acid metabolism (Groenen et al., 2004). Graça et al. performed several studies using different analytical techniques (NMR, UPLC-MS and MIR) to characterize a similar FM group comprising different types of malformations (central nervous system (CNS), abdominal, urogenital, cardiac and pulmonary), collected in 2^{nd} T (15-25 g.w). These studies unveiled similar results reinforcing a metabolic picture of fetal hypoxia, enhanced gluconeogenesis, TCA activity and hindered kidney development affecting FM pregnancies, newly noted changes were obtained by UPLC-MS reflecting lipid oxidation, altered placental amino acid transfer and changes in polyol pathways (Graça et al., 2009, 2010, 2012a, 2012b, 2013).

Regarding maternal plasma, an initial study addressed the possible role of nutritional intervention in preventing neural tube defects (NTD), revealing changes in 20 metabolites related to antioxidant capacity, proposed to relate to a reduction of NTD risk (Jiang et al., 2011). In a later study, GC-MS and UPLC-MS were used to characterize the sera metabolome of pregnant women (n=101) carrying NTD fetuses and controls (n=143), comprising all pregnancy trimesters (Zheng et al., 2011). Results unveiled 25 deregulated metabolites in NTD cases, suggesting a decreased activity of glycolysis, TCA cycle, mitochondrial electron transport, branched-chain amino acids and γ-aminobutyric acid metabolisms and methionine cycle. A parallel UPLC-MS study of 80 NTD affected women and 95 controls (mostly at < 28 g.w.) confirmed some of the above observations and suggested that, besides perturbation of the one-carbon unit metabolism, neurogenesis and oxidative stress are affected (Liang et al., 2012). The impact of fetal congenital heart defects (CHD) (n=27) on maternal serum metabolome

was also studied by MS and NMR, in 1st T (Bahado-Singh et al., 2014). The best markers for CHD prediction were hydroxypropionylcarnitine, glutaconylcarnitine and hydroxyl-tetradecadienylcarnitine (92.9% sensitivity and 93.2%, detected by direct infusion-MS), along with disturbances in lipid metabolism (PC and sphingolipids). In addition to the above studies, a general FM group (CNS, cardiac, urogenital, soft tissues and pulmonary malformations) was studied by ^{1}H NMR of 2nd T maternal plasma comprising results presented in this thesis (Diaz and Pinto et al., 2011).

Table 1.11. Metabolomics studies of fetal and newborn disorders by biological matrix. † lipidomics studies. BALF: bronchoalveolar lavage fluid, NB: newborn, RDS: Respiratory distress syndrome.

Matrix	Subject	References
Amniotic fluid	FM	(Bock, 1994), (Groenen et al., 2004), (Graça et al., 2009), (Graça et al., 2010), (Graça et al., 2012a), (Graça et al., 2012b), (Graça et al., 2013)
	CD	(Graça et al., 2010), (Graça et al., 2012a), (Amorini et al., 2012)
Placenta	FM	(Chi et al., 2014)
	SGA	(Horgan et al., 2010)
Maternal blood	FM	(Jiang et al., 2011), (Diaz and Pinto et al., 2011), (Zheng et al., 2011), (Graça et al., 2012a), (Liang et al., 2012), (Bahado-Singh et al., 2014)
	CD	(Diaz and Pinto et al., 2011), (Graça et al., 2012a), (Bahado-Singh et al., 2013b), (Bahado-Singh et al., 2013c), (Pinto et al., 2014d)†
	SGA	(Horgan et al., 2010), (Heazell et al., 2012)
	LBW & VLBW	(Ivorra et al., 2012), (Tea et al., 2012), (Alexandre-Gouabau et al., 2013)
	Macrosomia	(Ciborowski et al., 2014)
Maternal urine	FM	(Diaz and Pinto et al., 2011), (Graça et al., 2012a), (Graça et al., 2012b), (Diaz et al., 2013b)
	CD	(Diaz and Pinto et al., 2011), (Graça et al., 2012a), (Diaz et al., 2013b), (Trivedi and Iles, 2014b)
	IUGR	(Diaz et al., 2013b)
	Macrosomia	(Walsh et al., 2014)
Umbilical cord blood	SGA	(Horgan et al., 2011)
	LBW & VLBW	(Ivorra et al., 2012), (Tea et al., 2012), (Alexandre-Gouabau et al., 2013)
	IUGR	(Favretto et al., 2012), (Cosmi et al., 2013), (Sanz-Cortés et al., 2013)
	HIE	(Walsh et al., 2012), (Reinke et al., 2013)
NB blood	IEM	(Dénes et al., 2012), (Sahoo et al., 2012)
	Others	(Koulman et al., 2014)†, (Zhao et al., 2014)
NB urine	IUGR	(Dessì et al., 2011)
	Others	(Fanos et al., 2013c), (Fanos et al., 2014)
BALF	RDS	(Fabiano et al., 2011)
Hair	SGA	(Sulek et al., 2014)

A significant impact on maternal urine composition has been reported in FM cases also comprising heterogeneous groups (CNS, abdominal, urogenital, cardiac, limbs and poli-malformations) in the 2nd T by NMR and UPLC-MS (Diaz and Pinto et al., 2011, 2013b; Graça et al., 2012b). Results obtained by UPLC-MS for 22 FM and 26 control cases unveiled higher excretion of conjugated products suggesting alterations in gut microflora, glucuronic acid and sulphate conjugation reactions (Graça et al., 2012b). A more recent NMR study comprising higher number of samples of FM (n=35) and controls (n=84) revealed new changes being noted in enhanced gluconeogenesis, lipid oxidation and ketone body synthesis, respiratory chain and nucleotide metabolisms (Diaz et al., 2013b).

Chromosomal disorders (CD)

The impact of chromosomal disorders has been evaluated on amniotic fluid (Graça et al., 2010, 2012a; Amorini et al., 2012), maternal blood (Diaz and Pinto et al., 2011; Graça et al., 2012a; Bahado-Singh et al., 2013b, 2013c; Pinto et al., 2014d) and maternal urine (Diaz and Pinto et al., 2011, 2013b; Graça et al., 2012a). A small heterogeneous group of chromosomal disorders (n=10) was first studied by NMR, for which no relevant changes were found (Graça et al., 2010). In other study, the profile in amino group-containing compounds of amniotic fluid from 24 Down's syndrome-affected fetuses and 1257 normal pregnancies unveiled significant differences in Glu, Gln, Gly, Tau, Val, Ile, Leu, Orn, and Lys determined by HPLC (Amorini et al., 2012).

In relation to maternal blood, the first study addressed a general group of chromosomal disorders collected in the 2nd T of pregnancy and comprises results presented in this thesis (Diaz and Pinto et al., 2011). The second study (Bahado-Singh et al., 2013b) considered the 1st T (11-14 g.w.) NMR metabolic profile of sera of 30 women carrying T21 fetuses and 60 controls unveiling different levels of 11 metabolites, of which 3-hydroxybutyrate, 3-hydroxyisovalerate and 2-hydroxybutyrate were mostly responsible for group discrimination (75% sensitivity and 86.2% specificity). These results were interpreted as reflecting oxidative stress, poor myelination and neurotoxicity of the brain in T21 fetuses. Afterwards, Bahado-Singh et al. applied the same methodology to detect biomarkers of T18 in sera of 30 pregnant women compared to 114 controls and 30 T21 cases in the 1st T (11-14 g.w) (Bahado-Singh et al., 2013c). A combination of 2-hydroxybutyrate, glycerol, and maternal age gave 73.3% sensitivity and 96.6% specificity for T18 detection, while trimethylamine, Thr and creatine were sensitive for distinguishing T18 from T21. The possibility of an

increased antioxidant response in pregnancies with T18 was advanced by the authors. Finally, the most recent study of T21 cases using maternal blood samples was performed as part of this thesis with focus on plasma phospholipid profiling using a lipidomics MS approach (Pinto et al., 2014d).

The metabolic profile of maternal urine to unveil biomarkers of chromosomal disorders have been less explored including two reports considering a general CD group and T21 (Diaz and Pinto et al., 2011; Diaz et al., 2013b), and one specific to T21 (Trivedi and Iles, 2014b). The NMR-based urine profile of 33 women who had babies with general CD compared with 84 controls, at 2nd T (15-26 g.w.), resulted in a model with a sensitivity of 87% and a specificity of 84% which unveiled some effects similar to fetal malformations (other group studied in this paper) hinting at some sort of underlying general stress mechanism, while specific alterations included higher excretion of glucose and alterations in amino acid metabolism (Diaz et al., 2013b). In addition, the T21 group (n=13) was distinguished from other CD types (92% sensitivity and 95% specificity) by changes in energy metabolism (less excretion of glucose), as well as in nucleotide and amino acid metabolisms. In addition, the T21 group (n=13) was distinguished from other CD types by changes in energy metabolism (less excretion of glucose), as well as in nucleotide and amino acid metabolisms. Recently, a untargeted MS-based metabolomic study of urine of 23 women carrying a T21 fetus compared to 93 women with normal pregnancies between 9-23 g.w., revealed higher levels of dihydrouracil and lower levels of progesterone in the T21 group, along with changes in other metabolites not identified (m/z 109.1, 114.1 198.2 and 314.2) (Trivedi and Iles, 2014b). Changes in dihydrouracil were related with the liver abnormalities suffered by T21 fetuses during development, while progesterone may reflect impairment of placental function owing to oxidative stress.

Fetal growth anomalies (SGA, LBW, VLBW, IUGR, macrosomia)

Metabolomics studies of babies born small for gestational age (SGA) have been performed through the metabolic profiling of placental extracts (Horgan et al., 2010), and maternal and umbilical cord blood (Horgan et al., 2011; Heazell et al., 2012). These studies have been performed in small cohorts ($n \leq 9$) using only MS technologies (UPLC-MS). Placental extracts revealed alterations in 574 metabolite features related with SGA cases, suggesting that placental hypoxia and nutrient deficiency may play an important role in the pathophysiology of SGA (Horgan et al., 2010). Studies of maternal blood have been integrated with umbilical cord blood profiles in an attempt to obtain further insight into mother/fetus metabolic interplay.

Using maternal and cord plasma collected at 14-16 g.w., a set of 19 metabolites altered in SGA was found comprising several amino acids, carnitines, fatty acids, glycerophospholipids, hydroxyl acids, phosphocholines, phospholipids, sphingolipids, and steroid conjugates (Horgan et al., 2011). Comparison with maternal plasma composition supported the occurrence of placental dysfunction in SGA cases. In a later study (Heazell et al., 2012), 28 sera of SGA cases collected at 28-42 g.w. were considered together with other poor perinatal outcomes and 40 controls. Results unveiled decreases in glycerolipids, fatty and organic acids, glycerophospholipids, sterol lipids and vitamin D metabolites, and increased sphingolipids, in the poor outcome group.

Regarding LBW and VLBW studies, the biological matrices of choice to determine possible changes on maternal-fetal nutrient transfer and fetal metabolism have been the maternal and umbilical cord blood, analysed using both NMR and MS platforms. Ivorra et al. found a different NMR profile in umbilical cord blood of 20 LBW and 30 normal birth weight newborns, namely higher levels of Phe and citrulline, lower levels of Pro, choline, Gln, Ala, glucose, but no differences in maternal blood. Disturbances in placental amino acids transport, anabolic state and epigenetic regulation in LBW cases were suggested (Ivorra et al., 2012). Other similar comparative studies (sample cohorts ≤ 8) revealed alterations in materno-fetal nutrient transfer and fetal oxidative stress (Tea et al., 2012), and in fetal energy, antioxidant defence, and polyamines and purines flux (Alexandre-Gouabau et al., 2013) as signatures of preterm VLBW.

For IUGR cases, the first study performed to search for predictive biomarkers of IUGR in maternal urine was published by Diaz et al (Diaz et al., 2013b). Maternal urine was studied for 10 women who later gave birth to babies with IUGR (pre-IUGR) and 84 controls, collected at 15-26 g.w, suggesting main disturbances in amino acid profile indicated by a decrease of several amino acids (Ala, Gln, Gly, Leu, Lys, Thr and Tyr) determined by NMR spectroscopy. Other studies aimed at define the metabolic patterns of IUGR in the newborns using umbilical cord blood (Favretto et al., 2012; Sanz-Cortés et al., 2013) and newborn urine (Dessì et al., 2011). Untargeted liquid chromatography high-resolution MS (LC-HRMS) metabolic profiling of umbilical cord sera of 22 IUGR and 21 appropriate for gestational age (AGA) newborns showed differences in 22 metabolites, within which Phe, Trp and glutamate were the best IUGR discriminants (sensitivity 91-100%, specificity 85-89%) (Favretto et al., 2012). In addition, the NMR metabolic profile of umbilical cord plasma of 23 early IUGR vs. 23 matched AGA and 56 late IUGR vs. 56 matched AGA, with and without brain vasodilatation, showed an abnormal lipid metabolism in both cases mainly in unsaturated lipids of VLDL (Sanz-Cortés et al., 2013). Finally, the NMR metabolic

profiling of newborn urine of 26 neonates with IUGR and 30 neonates of proper gestational weight at birth, revealed differences in *myo*-inositol, sarcosine, creatine and creatinine (Dessì et al., 2011).

Macrosomia has been studied with the aim to search for predictive biomarkers in maternal serum (Ciborowski et al., 2014) and urine (Walsh et al., 2014). Maternal serum of women delivering normal birth weight newborns (n=28) and high birth weight (n=20), collected at 12-14 g.w., were analysed by LC-MS unveiling lower levels of several lipid classes (phospholipids, lysophospholipids and monoacylglycerols) and vitamin D3 metabolites, as well as increased levels of bilirubin in macrosomia cases (Ciborowski et al., 2014).

Other studies

Other applications of metabolomics in prenatal research include the studies respiratory distress syndrome (Fabiano et al., 2011), inborn errors of metabolism (IEM) (Dénes et al., 2012; Sahoo et al., 2012), hypoxic ischemic encephalopathy (HIE) (Walsh et al., 2012; Reinke et al., 2013), pesticide exposure (Bonvallot et al., 2013), delivery method (Hashimoto et al., 2013), cytomegalovirus (Fanos et al., 2013c), newborn's age (Koulman et al., 2014), biliary atresia and hepatitis syndrome (Zhao et al., 2014), breastfeeding (Sachse et al., 2014), and bronchopulmonary dysplasia (Fanos et al., 2014).

1.5. Aims of this thesis

This work has entailed five main goals:

1. To fill in existing gaps in the knowledge of human plasma degradability during handling and storage, and collection conditions such as fasting, paramount issues in NMR metabolomics. The contribution of this thesis to this issue was to address the effects of anti-coagulant collection tubes, room temperature stability, short- and long-term stability, freeze-thaw cycles and fasting vs. non-fasting state;
2. To define the metabolic alterations occurring in NMR plasma metabolome throughout healthy pregnancies, to unveil the metabolic trajectories of healthy pregnancies against to which deviant behaviours related to diseases may be identified;

3. To evaluate the prospects of maternal plasma metabolomics and lipidomics to unveil potential biomarkers (either individual metabolites or metabolite profiles) for the recognition of CD, in general, and T21, specifically. This study aims at contributing to the improvement of screening and diagnostic methods presently available in both 1st and 2nd trimesters;
4. To identify and interpret the metabolic alterations occurring in maternal plasma caused by GDM and to search for potentially predictive signatures of this condition before the onset of the disease;
5. To search for potential biomarkers of other prenatal disorders (FM, pre-PROM, pre-PTD and pre-preeclampsia) on 2nd trimester maternal plasma metabolome.

This thesis is organized as follows: chapter 1 describes the current status and challenges of prenatal health, the metabolomics strategy, and the state of the art of metabolomics in prenatal research; chapter 2 describes all experimental procedures performed; chapter 3 presents the plasma metabolome and lipidome of healthy subjects, as determined by NMR; chapter 4 evaluates the changes in the NMR plasma metabolic profile in connection to different handling and storage conditions, and with fasting/non-fasting conditions at collection; chapter 5 describes the metabolic trajectory found for maternal plasma throughout healthy pregnancies; chapter 6 and chapter 7 presents the results obtained for prenatal disorders with major impact in plasma metabolome and lipidome, namely CD (including T21) and GDM; chapter 8 includes the metabolic variations found for other prenatal disorders (FM, pre-PROM, pre-PTD and pre-preeclampsia) on NMR plasma metabolome; finally chapter 9 presents the general conclusions and future perspectives. The results presented in chapters 4 to 7 were fully published or submitted in peer-reviewed journals, while only part of the results of chapter 8 were published. All published and submitted articles are here adapted slightly onto the form of a thesis chapter, for the sake of clarity.

Chapter 2. Experimental procedures

The present chapter describes the sample and metadata collection procedures, as well as the experimental and data analysis strategies used to conduct the experimental work presented throughout the thesis. This work was carried out with the collaboration of the Maternity Bissaya Barreto (MBB), Centro Hospitalar de Coimbra, where all samples were collected, from 2008 to 2013, and the Faculty of Medicine of the University of Coimbra (FMUC) where samples were stored immediately after collection. Between January 2008 and June 2011, this work was supported by the Portuguese Foundation of Science and Technology (FCT) funding through the PTDC/QUI/66523/2006 project. Samples were collected under ethical committee approval (Ref. 18/04 and 29/096) and informed consents were obtained, from all subjects participating in the study (see Annex for informed consents and individual questionnaires). The donors comprised Caucasian pregnant women from the central part of Portugal.

2.1. Sampling

2.1.1. Samples and clinical metadata

For pregnant women, plasma samples were collected during medical appointments, in the morning period, under non-fasting conditions due to restrictions in controlling/limiting pregnant women's diet. For non-pregnant women, samples were collected in fasting (overnight) conditions and non-fasting during the morning period (2h after breakfast). Whole blood was collected into sodium heparin tubes (Vacuette® Tube Sodium Heparin 9 mL) and centrifuged (1500 *g*, 10 min, 4°C) within 30 min of collection as illustrated in Figure 2.1. Supernatants were frozen at -20ºC for up to 2 hours in MBB. Then, all samples were transferred from MBB to FMUC, where they were given a random code (to ensure patients confidentiality) and stored at -80ºC. Samples were then transported, in dry ice, to CICECO-Chemistry Department, at the University of Aveiro (UA) where they were stored and analysed.

Clinical and obstetrical metadata including pre-pregnancy and prenatal health history were obtained through individual questionnaires filled at the time of collection and at the end of gestation for all pregnancies. Relevant parameters included pre-existing diseases, medications, lifestyle and nutritional habits, blood type, maternal height and weight before pregnancy (used to calculate the body mass index, BMI), smoking habits, obstetrical background, blood pressure records, blood chemistry and urine analysis results, fetal karyotype and ultrasound monitoring results. Information about labour (or pregnancy termination for cases of FM or CD with poor prognostic)

and neonatal outcome was also obtained for each sample. The gathered metadata allowed the classification of the samples into controls and several disorder groups. Samples were classified into controls, if they were obtained from healthy pregnant women with a normal pregnancy course and delivered healthy term newborns, no pre-existing maternal pathologies and not under medication. Criteria for the definitions of other groups are described throughout this section.

Overall from 2008 to 2013, 942 blood plasma samples were collected, out of which 297 were found eligible for this work. For the remaining samples, information obtained after labour revealed pre-existing diseases (e.g. chronic hypertension, diabetes mellitus, thyroid disorders), infections (e.g. cytomegalovirus), medication and incomplete metadata information (due to delivery in other hospitals).

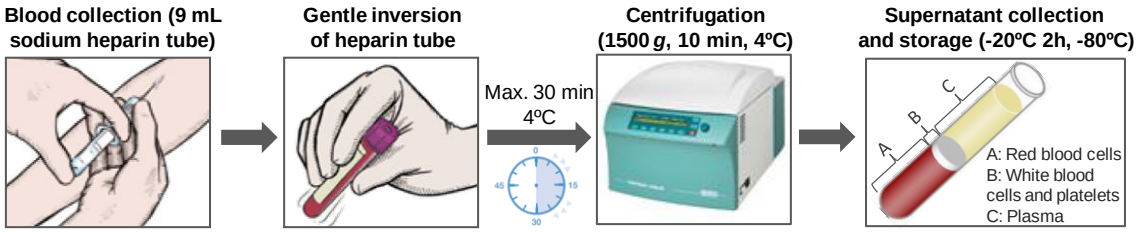


Figure 2.1. Schematic representation of the blood plasma collection protocol carried out in this thesis.

2.1.2. Definition of sample groups

This thesis is divided into five main studies described in Chapters 4 to 8. In Study 1 (Chapter 4), key points of plasma stability during handling and storage were studied in order to fill some existing gaps in the knowledge about human plasma degradability. Study 2 (Chapter 5) aimed the identification of metabolic adaptations reflected in maternal blood plasma throughout healthy pregnancies and correlation to urine. In Study 3 (Chapter 6), the possibility of using 1st and 2nd T maternal blood plasma for the diagnosis of CD and T21 was explored. Study 4 (Chapter 7) aimed the identification and interpretation of blood plasma metabolic alterations related with pre and post-diagnosis GDM cases in the 2nd T. Finally in Study 5 (Chapter 8), the effects of other prenatal disorders on maternal plasma metabolome of 2nd T were studied, namely FM and pre-diagnosis PROM, preterm delivery and preeclampsia. A schematic representation of Studies 1 to 5 with corresponding aims is shown in Figure 2.2.

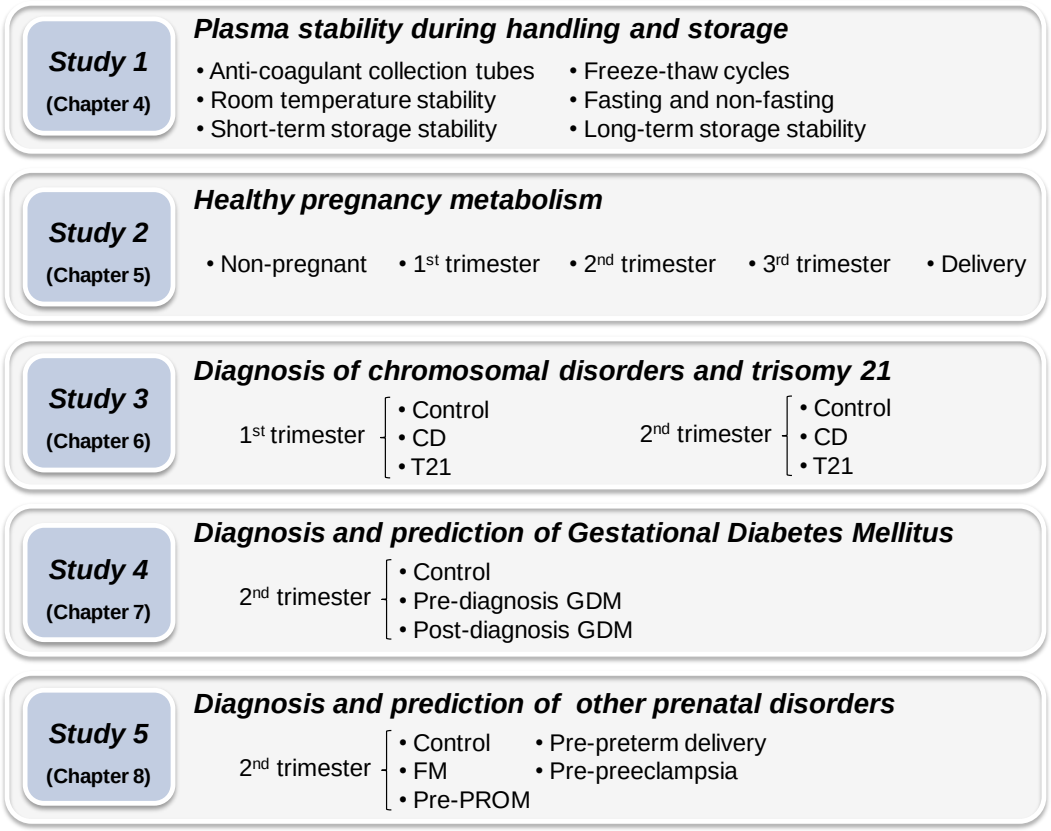


Figure 2.2. Schematic representation of Studies 1 to 5 with corresponding aims.

Sample groups in study 1: Evaluation of plasma stability during handling and storage by NMR metabolomics

Table 2.1 lists the groups of subjects employed in this work, which include, for some studies, healthy pregnant women in their 2nd T of pregnancy. Such samples do exhibit few compositional differences, compared to those from non-pregnant subjects (in particular, higher lipid levels), however, their stability properties are expected to be at least qualitatively equivalent.

In order to study the effect of heparin and ethylene diamine tetraacetic acid (EDTA) tubes on plasma metabolome, whole blood of 5 non-fasting subjects was collected for both an EDTA tube (Vacuette® Tube EDTA K2 6mL) and a sodium heparin tube (Vacuette® Tube Sodium Heparin 9 mL). Supernatants were frozen at -20ºC for up to 2 hours and then stored at -80ºC (3 days) until NMR analysis. For all other studies, blood plasma samples were collected into sodium heparin tubes. To evaluate the effects of storage at room temperature, 3 samples were thawed, prepared for NMR analysis and kept in the probe for 21 h. For the short-term stability study (at -20^{o}C and -80^{o}C),

samples were split into aliquots prior to freezing, each sample being thawed after 3, 7 and 31 days. For the freeze-thaw cycles study, each cycle involved sample thawing at room temperature for 30 min, followed by refreezing at -80°C. For the comparison of the NMR-metabolic profile of fasting and non-fasting subjects, plasma was collected from the same 16 subjects under fasting (range 8-14h) and non-fasting conditions (2 h later, after non-standardised breakfast). In order to determine the impact of long-term storage on plasma metabolome, plasma of different subjects were divided into three groups according the length of storage at -80ºC: 0.5 to 0.9 year (n=11), 1.2 to 1.6 years (n=17) and 1.7 to 2.5 years (n=21).

Table 2.1. Blood plasma sample groups and corresponding identification of studies, number of subjects (*n*), age and BMI. In columns age and BMI, numbers in brackets indicate average values, for groups with larger number of subjects. The last column indicates that, for some studies, plasma samples were collected from healthy pregnant women in their 2nd T of pregnancy (between 16 and 24 g.w.).

Study	*n*	Age /years	BMI /Kg.m^{-2}	Observations
Anti-coagulant collection tubes	5	23-49	19.8-34.8	Non-pregnant
Room temperature stability	3	29-37	21.8-32.5	Pregnant (16 g.w.)
Short-term -20°C and -80°C storage stability	3	25-35	19.8-34.8	Non-pregnant
Freeze-thaw cycles	3	33-36	20.9-26.4	Pregnant (17 g.w.)
Fasting and non-fasting	16	21-36 (27.4)	17.8-26.6 (22.0)	Non-pregnant
Long-term -80°C storage				
6-12 months	11	25-41 (34.1)	19.8-27.7 (23.6)	Pregnant (16-22 g.w.)
14-19 months	17	28-42 (35.7)	20.4-25.6 (23.3)	Pregnant (16-20 g.w.)
20-30 months	21	31-42 (36.3)	20.3-33.0 (24.0)	Pregnant (16-24 g.w.)

Sample groups in study 2: Following healthy pregnancy metabolism by NMR metabolomics of blood plasma

In order to evaluate the metabolic adaptations reflected in plasma throughout healthy pregnancies, blood plasma samples were collected for healthy non-pregnant (NP) and pregnant women in their 1st, 2nd and 3rd trimesters of pregnancy, at their routine medical appointments, and at the time of delivery. All four groups were

independent (due to the logistics of sample collection in the clinic), therefore providing an average description of each pregnancy trimester. Consideration of representative populations may present advantages over individual follow-up since results are not dependent on individual phenotypes and identifies the more marked effects of pregnancy. Table 2.2 lists sample numbers and characteristics of each group of subjects considered, a lower number of subjects (n=7) being obtained for the delivery group.

Table 2.2. List of whole blood plasma samples and lipid extracts collected for each independent group of subjects along with corresponding number of samples (*n*), ranges of maternal age (in years), gestational age at sampling (in gestational weeks) and pre-pregnancy body mass index (BMI, in $kg.m^{-2}$). Values in brackets correspond to average values. T: trimester.

Subjects group	*n*	Maternal age /years	Gestational age /weeks	Pre-pregnancy BMI /$kg.m^{-2}$
Whole blood plasma				
Non-pregnant	20	21-48 (28)		18-32 (23) [for n=20]
1st T	25	19-37 (30)	11-13 (12)	19-38 (26) [for n=18]
2nd T	30	25-42 (36)	16-22 (17)	20-33 (24) [for n=25]
3rd T	12	22-38 (31)	31-40 (37)	21-28 (23) [for n=9]
Delivery	7	26-38 (32)	32-41 (38)	20-25 (23) [for n=3]

Sample groups in study 3: First and second trimester maternal plasma for the diagnosis of chromosomal disorders and trisomy 21

Maternal blood plasma was collected for women carrying fetuses diagnosed with CD and for healthy pregnant women (controls) in their 1st and 2nd T. These groups were independent since CD diagnosis was performed in one of the trimesters only. 1st T and 2nd T CD cases were diagnosed by CVS (12-13 g.w.) and amniocentesis (14-23 g.w.), respectively. Sample numbers and group details are listed in Table 2.3, according to sub studies considering whole plasma profile, plasma lipid profile and plasma phospholipid profile. For plasma phospholipid profiling, plasma pools of different subjects were used thus enabling inter-individual variability to be averaged out and average T21 features to be identified.

Table 2.3. List of maternal samples collected for each group of subjects in *Study 3*, along with corresponding maternal age/years and gestational age/weeks at the time of collection, and pre-pregnancy BMI/kg.m^{-2}. Values in brackets correspond to average values. *n*: number of samples; CD: chromosomal disorders; T21: trisomy 21; T: trimester.

Subjects group	*n*	Maternal age /years	Gestational age /weeks	Pre-pregnancy BMI / kg.m^{-2}
Whole blood plasma				
Controls 1st T	25	24-37 (31)	11-13 (13)	19-38 (25) [for n=23]
CD 1st T	15	21-44 (35)	12-13 (13)	19-39 (26) [for n=9]
T21 1st T	8	21-44 (35)	12-13 (13)	21-30 (24) [for n=5]
Controls 2nd T	49	25-42 (36)	16-24 (17)	20-33 (24) [for n=37]
CD 2nd T	30	25-43 (36)	14-23 (17)	19-34 (26) [for n=26]
T21 2nd T	14	26-43 (36)	15-23 (17)	20-33 (25) [for n=12]
Plasma lipid extracts				
Controls 1st T	15	24-37 (31)	12-13 (13)	22-35 (25) [for n=15]
T21 1st T	7	32-44 (37)	12-13 (13)	21-30 (25) [for n=4]
Controls 2nd T	15	28-42 (37)	16-22 (17)	20-26 (22) [for n=11]
T21 2nd T	12	26-43 (37)	14-23 (17)	22-33 (27) [for n=8]
Plasma phospholipid extracts (pools)				
Control pool 1 2nd T	5	30-38 (34)	16-18 (17)	26 [for n=1]
Control pool 2 2nd T	4	34-39 (36)	16-17 (17)	21-24 (22) [for n=4]
Control pool 3 2nd T	5	31-39 (36)	16-17 (17)	20-26 (23) [for n=4]
T21 pool 2nd T	5	39-43 (41)	15-18 (17)	22-33 (28) [for n=3]

The CD group comprise cases from diverse types of CD in 1st and 2nd T, listed in Table 2.4.

Table 2.4. List of specific types of CD along with corresponding number of samples (*n*). T: trimester.

Chromosomal disorders (CD) 1st T		Chromosomal disorders (CD) 2nd T	
Type	***n***	**Type**	***n***
Trisomy 21	8	Trysomy 21	14
Trisomy 13	2	Trisomy 18	2
Trisomy 9	1	Kleinfelter syndrome	2
Triploidy (69, XXX)	1	XYY syndrome	2
Translocations (familial)	2	XXX syndrome	1
Mosaicism (autosomal)	1	Triploidy (69, XXX)	1
		Mosaicism (autosomal)	1
		Marker chromosome (autosomal)	1
		Translocations (2 familial + 1 de novo)	3
		Inversions (familial)	2
		Deletion (autosomal)	1

Sample groups in study 4: Second trimester maternal plasma for the diagnosis and prediction of gestational diabetes mellitus (GDM) by NMR metabolomics

For the study of metabolic alterations in 2nd T maternal plasma of pregnancies affected by GDM, samples were collected from independent groups of pregnant women during routine medical appointments and diabetes appointments. Blood samples were collected for healthy pregnant women, pregnant women who later developed GDM (pre-diagnosis GDM group, 2-21 g.w. prior to diagnosis) and women after GDM diagnosis (post-diagnosis GDM group, 1-2 weeks after the OGTT test, at 18-38 g.w., and before any clinically advised treatment). GDM was diagnosed after a 75g or 100g oral OGTT by measuring plasma glucose at fasting and after 1, 2 and 3 (for the latter) hours according to the International Association of Diabetes and Pregnancy Study Groups criteria (Metzger et al., 2010). Sample numbers, maternal age range, gestational age range at collection, interval of time from sampling to diagnosis, gestational age range at delivery, baby weight and pre-pregnancy BMI range can be found in Table 2.5.

Table 2.5. List of maternal plasma samples and lipid extracts obtained for each group of subjects, with corresponding number of samples (n); maternal age (years); gestational age (weeks) at collection and, for the pre-diagnosis GDM group, from collection to diagnosis; gestational age (weeks) at delivery; baby weight at birth (g) and pre-pregnancy body mass index (BMI, $kg.m^{-2}$). Values in brackets correspond to average values. [a] Pre-diagnosis GDM includes cases of hypertension (n=1), twin pregnancy (n=2), preterm delivery (n=3), macrosomia (n=1) and intrauterine growth restriction (n=1). [b] Post-diagnosis GDM includes cases of hypertension (n=1), twin pregnancy (n=2) and preterm delivery (n=1). [c] This group includes one case of hypertension.

Subjects group	*n*	Maternal age /years	g.w. at collection /weeks	g.w. to diagnosis /weeks	g.w. at delivery /weeks	Baby weight (g)	Pre-pregnancy BMI /$kg.m^{-2}$
Whole blood plasma							
Controls	49	25-42 (36)	16-24 (17)	-	37-41 (39)	2405-3960 (3222)	20-33 (24) [for n=37]
Pre-diagn. GDM[a]	32	30-44 (38)	16-21 (17)	2-21 (13)	35-40 (38)	2125-3805 (3115)	18-35 (26) [for n=24]
Post-diagn. GDM[b]	12	18-41 (33)	24-27 (26)	-	34-40 (38)	1745-3505 (2736)	18-36 (26) [for n=12]
Plasma lipid extracts							
Controls	15	28-42 (37)	16-22 (17)	-	38-40 (39)	2560-3960 (3288)	20-26 (22) [for n=11]
Pre-diagn. GDM[c]	14	36-42 (39)	16-19 (17)	2-19 (12)	36-40 (38)	2310-3805 (3109)	18-35 (24) [for n=14]
Post-diagn. GDM[b]	12	18-41 (33)	24-27 (26)	-	34-40 (38)	1745-3505 (2756)	18-36 (26) [for n=12]

Sample groups in study 5: Study of the effects of other prenatal disorders on second trimester maternal plasma metabolome

Maternal blood plasma was collected for women carrying fetuses diagnosed with fetal malformations (FM) and for other women with no apparent disorder at the time of collection developing later, premature rupture of the membranes (Pre-PROM), preterm delivery (pre-preterm delivery) and preeclampsia (pre-preeclampsia). For the pre-preterm and pre-preeclampsia groups, a preliminary study of plasma phospholipids was performed considering plasma pools of different subjects. Sample numbers, maternal age, gestational ages at sampling, pre-pregnancy BMI and interval of time from sampling to diagnosis for pre-diagnostic disorders, are listed in Table 2.6.

Table 2.6. List of maternal plasma samples and plasma phospholipid extracts collected for each group of subjects in *Study 5*, along with corresponding number of samples (*n*), maternal age/years and gestational age/weeks at the time of collection, pre-pregnancy BMI/kg.m^{-2}, and for the pre-diagnosed group, g.w. between sampling and time of diagnosis. Values in brackets correspond to average values.

Subjects group	*n*	Maternal age /years	Gestational age /weeks	Pre-pregnancy BMI /kg.m^{-2}	g.w. to diagnosis
Whole blood plasma					
Controls	49	25-42 (36)	16-24 (17)	20-33 (24) [for n=37]	-
FM	36	19-45 (30)	14-27 (20)	18-36 (25) [for n=25]	-
Pre-PROM	36	27-44 (37)	16-20 (17)	20-33 (24) [for n=26]	19-24 (22)
Plasma phospholipid extracts (pools)					
Control 1	5	30-38 (34)	16-18 (17)	26 [for n=1]	-
Control 2	4	34-39 (36)	16-17 (17)	21-24 (22) [for n=4]	-
Control 3	5	31-39 (36)	16-17 (17)	20-26 (23) [for n=4]	-
Pre-preterm	5	34-38 (37)	17-19 (17)	23-25 (24) [for n=4]	14-18 (17)
Pre-preeclampsia	4	35-41 (37)	16-17 (17)	19-33 [for n=3]	14-18 (16)

FM cases were diagnosed by morphological ultrasound after 12 g.w., the group comprising cases of central nervous system (CNS), cardiac, limbs, urogenital, polimalformations and others, as listed in Table 2.7. Pre-PROM include women who had rupture premature of the membranes at 37-41 g.w., pre-preterm cases include women who had preterm deliveries at 29-36 g.w., and pre-preeclampsia comprise

women diagnosed at 31-34 g.w. through blood pressure ≥ 140/90 mmHg and proteinuria ≥ 0.3 g/L in 24-h urine. No sufficient number of samples were gathered for other relevant disorders such as intrauterine growth restriction (n=4), fetal death (n=5), small for gestational age (n=4), and large for gestational age (n=6).

Table 2.7. List of specific types of FM along with corresponding number of samples (*n*).

Fetal malformations (FM) 2nd T			
Type	***n***	**Type**	***n***
CNS	**13**	**Polimalformation**	**6**
Corpus callosum agenesis	3	Club feet + left pielectasy + choroid plexus cysts	1
Arnold-Chiari	2		
Holoprosencephaly	2	Cystic hygroma + bladder agenesis	1
Ventriculomegaly	2		
Ventriculomegaly + polydactyly	1	Corpus callosum agenesis + pancreatic agenesis + bile ducts malformation	1
Spina bifida	1		
Craniosynostosis	1		
Myelomeningocele	1	Cystic hygroma + jugular sacs	1
Cardiac	**7**	Diffuse renal dysplasia + adenomatoid malformation of the right lung	1
Cardiopathy	5		
Hypoplastic left heart syndrome	1		
Dextrocardia	1	Severe bilateral pleural effusion + severe cervical edema + club feet + overlapping fingers in hands	1
Limbs	**4**		
Club feet	2		
Polydactyly	1	**Others**	**4**
Lower limbs malformation	1	Enlarged stomach	1
Urogenital	**3**	Cystic adenomatoid malformation in the right lung	1
Hydronephrosis	1		
Bladder agenesis	1	Congenital diaphragmatic hernia	1
Hydronephrosis + pielectasy	1	Congenital disorder of glycosylation type Ia	1

2.2. Chemicals

All solvents and solid chemicals used in the experimental work were of HPLC grade or equivalent and used as received, a) deuterated solvents: chloroform-d ($CDCl_3$), $CDCl_3$ with 0.03% tetramethylsilane (TMS) and deuterium oxide (D_2O) (Eurisotop, France); b) non-deuterated solvents: acetonitrile (Fisher Scientific, Portugal), ammonium acetate (Merck Millipore, VWR, Portugal), ammonium hydroxide, formic acid and hexane (Fluka, Sigma Aldrich, Portugal), chloroform ($CHCl_3$) (HiPerSolv Chromanorm, VWR, Portugal), Isopropanol (2-propanol) (Normapur, VWR, Portugal), Methyl-tert-butyl ether (MTBE) (Chromasolv, Sigma Aldrich, Portugal); and c) solid chemicals: 1,2-dimyristoyl-sn-glycero-3-phosphocholine (14:0/14:0 PC), 1-palmitoyl-2-

hydroxy-sn-glycero-3-phosphocholine (16:0 LPC), 1-palmitoyl-2-linoleoyl-sn-glycero-3-phosphocholine (16:0/18:2 PC), 1-palmitoyl-2-linoleoyl-sn-glycero-3-phosphoethanolamine (16:0/18:2 PE) and N-palmitoyl-D-erythro-sphingosylphos-phorylcholine (d18:1/16:0 SM) (Avanti Polar Lipids, Instruchemie, The Netherlands), cholesteryl octadecadienoate (18:2 CE), cholesteryl eicosatetraenoate (20:4 CE) and sodium azide (Sigma Aldrich, Portugal), sodium chloride (BioXtra, Sigma Aldrich Portugal).

2.3. Profiling of maternal plasma by NMR spectroscopy

2.3.1. Sample preparation

Whole blood plasma

Plasma samples were thawed at room temperature and 400 µL of saline solution (NaCl 0.9% in 10% D_2O, with 3mM NaN_3 for part of the room temperature study in Study 1 (Chapter 4)) were added to 200 µL of plasma. The mixture was then centrifuged (4500 *g*, 25ºC, 5 min) and 550 µL transferred into 5 mm NMR tubes.

Plasma lipid extracts

In order to study the NMR plasma lipidome, lipids from plasma samples were extracted in duplicate according to the methyl-tert-buthyl ether (MTBE) method (Matyash et al., 2008). This method was previously compared with other lipid extraction methods (Folch, Bligh and Dyer, acidified Bligh and Dyer, and hexane-isopropanol), in the context of a master thesis (Almeida, 2014), unveiling better reproducibility and easier protocol. A schematic representation of the MTBE-methanol extraction protocol is shown in Figure 2.3. Firstly, ice-cold (-20ºC) methanol (MeOH, 1.5 mL) was added to blood plasma (200 µL) and the mixture vortexed (60 s), followed by the addition of ice-cold MTBE (5.0 mL) and vortex mixing (60 s). Then, the mixture was incubated for 1 hour on ice, with occasional vortex mixing. Phase separation was induced by adding ultrapure water (ultrapure H_2O, 1.25 mL) and let it stand for 10 min on ice, with occasional mixing. After centrifugation (1000 *g*, 10 min), the organic phase (upper) was collected and the lower phase was reextracted with 2 mL of a mixture of ice-cold MTBE/MeOH/UP-H_2O (10:3:2.5, v/v/v). Combined organic phases were dried in a vacuum centrifuge, followed by storage at -80ºC until NMR analysis.

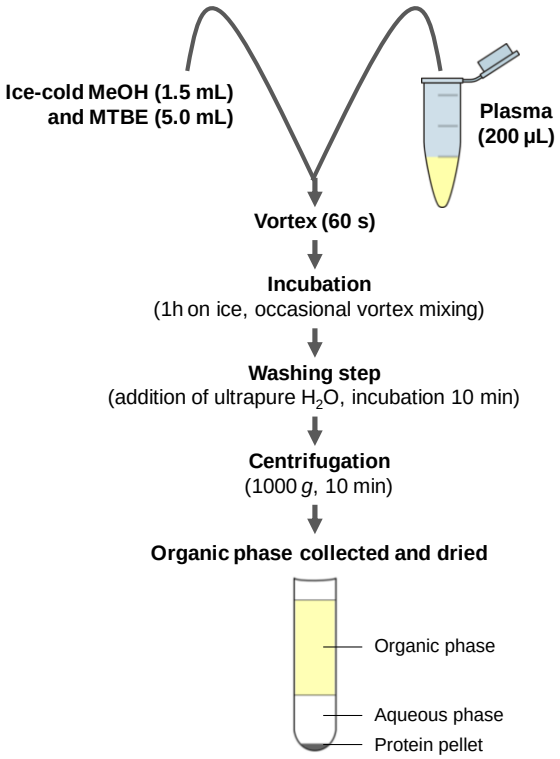


Figure 2.3. General plasma lipid extraction steps performed using the methanol-MTBE method.

Before NMR analysis, lipid extracts were thawed at room temperature, followed by addition of 600 µL of chloroform-d with tetramethylsilane (TMS) (99.96% + 0.03%, v/v). The mixture (580 µL) was then transferred into a 5 mm NMR tube.

Details on urine collection and NMR analysis for Study 2 and Study 3 can be found elsewhere (Diaz et al., 2013a, 2013b).

2.3.2. NMR analysis

1D and 2D NMR acquisition of whole blood plasma

The 1D and 2D NMR spectra were recorded, at 300 K, on a Bruker Avance DRX 500 spectrometer operating at a 500.13 MHz frequency for proton and equipped with an actively shielded gradient unit with a maximum gradient strength output of 53.5 G/cm. For each blood plasma sample, three 1D ^{1}H NMR spectra were obtained: a standard spectrum (1D NOESY-presat), a Carr-Purcell-Meiboom-Gill (CPMG) spectrum and a diffusion-edited spectrum, using the noesypr1d (RD-90°-t-90°-t_m-90°-ACQ), cpmgpr (RD-90°-{τ-180°-τ}$_n$-ACQ) (Meiboom and Gill, 1958) and ledbpgp2s1dpr (RD-90°-G1-180°-G1-90°-G2-T-90°-G1-180°-G1-90°-G2-τ-90°-acquire FID) (Wu et al.,

1995) pulse sequences, respectively, from the Bruker library. Each free induction decay (FID) was zero-filled to 64 k points and multiplied by a 0.3 Hz exponential line-broadening function prior to Fourier transformation. Spectra were manually phased and baseline corrected and chemical shifts referenced internally to the α-glucose H1 resonance (at 5.23 ppm). A list of all acquisition and processing parameters is show in Table 2.8.

Table 2.8. List of acquisition and processing parameters used for 1D and 2D NMR experiments of blood plasma. NS: number of scans, RD: relaxation delay, t_m: mixing time, *t*: fixed delay, τ, echo time; n, number of (τ-180°-τ) loops; 2nτ, total spin-spin relaxation time; δ, gradient duration; Δ, diffusion time; g, strength gradient pulses as % of maximum gradient output (53.5 G/cm); SW: spectral window, TD: size of FID, RG: receiver gain, SI: size of real spectrum, WDW: type of window function, qsine: sine-bell-squared function, F1: 1st dimension, F2: 2nd dimension, LB: line broadening,

Experiment	Acquisition parameters	Processing parameters
Standard 1D (noesypr1d)	NS: 128; RD: 4 s; t_m: 100 ms; t: 3 µs; SW: 10330.58 Hz/ 20.66 ppm; TD: 32 k; RG: 64	SI: 64 k; WDW: exponential; LB: 0.30 Hz; Manual phase and baseline correction;
CPMG (cpmgpr)	NS: 256; RD: 4 s; τ: 400 µs; n: 80; 2nτ: 64 ms; SW: 10330.58 Hz/ 20.66 ppm; TD: 32 k; RG: 128	Chemical shift calibration by α-glucose H1 resonance (at 5.23 ppm)
Diffusion-edited (ledbpgp2s1dpr)	NS: 128; RD: 4 s; δ: 2 ms; Δ: 100 ms; g: 90 % ; SW: 10330.58 Hz/ 20.66 ppm; TD: 32 k; RG: 32	
TOCSY (clmlevprtp)	NS: 32; RD: 2 s; t_m: 96 ms; pl10: 15 dB; TD [F1]: 128; TD [F2]: 2 k; SW [F1]: 8012.82 Hz; SW [F2]: 8012.82 Hz	SI [F1]: 1 k; SI [F2]: 4 k; WDW: qsine; Manual phase correction; Chemical shift calibration by α-glucose H1 resonance (at 5.23 ppm)
HSQC (invietgpsi)	NS: 48; RD: 2 s; TD [F1]: 300; TD [F2]: 4 k; SW [F1]: 25153.81 Hz; SW [F2]: 8012.82 Hz	
JRES (jresgpprqf)	NS: 16; RD: 2 s; TD [F1]: 64; TD [F2]: 8 k; SW [F1]: 43.41 Hz; SW [F2]: 8012.82 Hz	

The TOCSY experiment was recorded in phase-sensitive mode using States-TPPI (time proportional phase incrementation) detection in t_1 using a MLEV-17 pulse sequence (Bax and Davis, 1985), clmlevprtp. The HSQC spectra were recorded with inverse detection and ^{13}C decoupling during acquisition using the invietgpsi pulse program. The JRES experiment was acquired using the jresgpprqf pulse sequence with presaturation during relaxation delay. The acquired FIDs were zero-filled according the data points defined in SI parameters shown in Table 2.8 and multiplied by a sine-bell-squared function (qsine) prior to FT. All spectra were manually phased and baseline corrected, and chemical shifts calibrated internally using α-glucose H1 resonance (at

5.23 ppm). All NMR data processing was carried out in TOPSPIN 2.0 and 3.2 versions (Bruker BioSpin, Reinstetten, Germany).

Peak assignments were carried out using the 2D NMR experiments (TOCSY, HSQC and JRES), the Bruker Biorefcode spectral database, the Biological Magnetic Resonance Bank (BMRB) and the Human Metabolome Database (HMDB) (Seavey et al., 1991; Wishart et al., 2007).

1D and 2D NMR acquisition of plasma total lipid extracts

NMR spectra were recorded, at 298 K, on a Bruker Avance DRX 500 spectrometer operating at a 500.13 MHz frequency for proton and equipped with a triple resonance inverse (TXI) probe head. A list of all acquisition and processing parameters is show in Table 2.9. 1D NMR spectra were acquired using a standard Bruker proton pulse sequence (zg) with a 90º pulse, the acquisition time was 2.34 s with a relaxation delay of 5 s. Each FID was zero-filled to 64 K points, multiplied by a 0.3 Hz exponential line-broadening function prior to Fourier transform. Spectra were manually phased and baseline corrected and chemical shifts referenced internally to TMS (at 0.0 ppm).

Table 2.9. List of acquisition and processing parameters used for 1D and 2D NMR experiments of plasma total lipid extracts. NS: number of scans, RD: relaxation delay, t_m: mixing time, SW: spectral window, TD: size of FID, RG: receiver gain, SI: size of real spectrum, WDW: type of window function, qsine: sine-bell-squared function, F1: 1st dimension, F2: 2nd dimension, LB: line broadening.

Experiment	Acquisition parameters	Processing parameters
Standard ^{1}H (zg)	NS: 128; RD: 5 s; SW: 7002.80 Hz/ 14.00 ppm; TD: 32 k; RG: 32	SI: 64 k; WDW: exponential; LB: 0.30 Hz; Manual phase and baseline correction; Chemical shift calibration by TMS (0.0 ppm)
TOCSY (dipsi2ph)	NS: 16; RD: 2 s; t_m: 70 ms; pl10: 15 dB; TD [F1]: 256; TD [F2]: 4 k; SW [F1]: 7002.80 Hz; SW [F2]: 7002.80 Hz	SI [F1]: 1 k; SI [F2]: 4 k; WDW: qsine; Manual phase correction; Chemical shift calibration by TMS (0.0 ppm)
HSQC (hsqcetgp)	NS: 32; RD: 2 s; TD [F1]: 128; TD [F2]: 2 k; SW [F1]: 20831.98 Hz; SW [F2]: 7002.80 Hz	
J-resolved (jresgpprqf)	NS: 16; RD: 2 s; TD [F1]: 40; TD [F2]: 8 k; SW [F1]: 43.41 Hz; SW [F2]: 7002.80 Hz	

2D NMR experiments were acquired for assignment namely TOCSY, HSQC and *J*-res. TOCSY spectra were acquired in phase-sensitive mode using States-TPPI

detection with a DIPSI 2 sequence for mixing using the dipsi2ph pulse program. ^{1}H-^{13}C HSQC spectra were acquired using the hsqcetgp program with $^{1}H/^{13}C$ correlation via double inept transfer in phase-sensitive mode using Echo/Antiecho-TPPI gradient selection with decoupling during acquisition and using trim pulses in inept transfer. JRES spectra were acquired using the jresgpprqf pulse program with presaturation during relaxation delay and using gradients. The acquired FIDs were zero-filled according the data points defined in SI parameters shown in Table 2.9 and multiplied by a sine-bell-squared function (qsine) prior to FT. All spectra were manually phased, automatically baseline corrected and manually chemical shifts calibrated internally using TMS (at 0.0 ppm). All NMR data processing was carried out in TOPSPIN 3.2 (Bruker BioSpin, Reinstetten, Germany).

Peak assignments were carried out using the above described 2D NMR experiments (TOCSY, HSQC and JRES), as well as lipid standards and the multiple-quantum NMR experiments described below.

Multiple quantum (MQ) NMR acquisition for assignment of lipid extracts

For further assignment and identification of overlapping signals in proton spectra of plasma lipid extracts, the potential of multiple quantum (MQ) NMR spectroscopy was evaluated using a simplified lipid mixture composed by a set of four phospholipids and two cholesteryl esters. Phospholipid standards included 1-palmitoyl-2-hydroxy-sn-glycero-3-phosphocholine (16:0 LPC), 1-palmitoyl-2-linoleoyl-sn-glycero-3-phospho-choline (16:0/18:2 PC), 1-palmitoyl-2-linoleoyl-sn-glycero-3-phosphoethanolamine (16:0/18:2 PE) and N-palmitoyl-D-erythro-sphingosylphosphorylcholine (d18:1/16:0). Cholesteryl ester standards included cholesteryl octadecadienoate (18:2 CE) and cholesteryl eicosatetraenoate (20:4 CE). All lipid standards were analysed in pure state (1.25 mM in $CDCl_3$) and in a mixture (1 mM of each one in $CDCl_3$).

NMR experiments were performed by courtesy of Institut des Sciences Moléculaires de Marseille (UMR 7313 iSm2), Aix Marseille Université. NMR acquisition was performed at 298 K using a Bruker Avance-600 NMR spectrometer equipped with a TXI probe capable of generating gradients fields up to 60 G/cm. First, a set of 1D and 2D NMR experiments were performed, namely a standard ^{1}H spectrum (zg) and ^{1}H-^{1}H correlation spectroscopy (COSY). The COSY experiment was performed using a 2D homonuclear shift correlation with gradient pulses for selection and purge pulses before relaxation delay, with the cosygpppqf pulse sequence from Bruker library. A list of all acquisition and processing parameters is show in Table 2.10.

Table 2.10. List of acquisition and processing parameters used for 1D and 2D NMR experiments and multiple quantum NMR experiments of lipid standards and lipid mixture. NS: number of scans, RD: relaxation delay, Q: quantum, τ: delay for MQ creation, t_1: fixed delay, t_2: delay for gradient recovery, G_1: first pulsed field gradient as % of maximum gradient output (60 G/cm),G_2: second pulsed field gradient as % of maximum gradient output (60 G/cm), SW: spectral window, F1: 1st dimension, F2: 2nd dimension, O1: transmitter frequency offset, TD: size of FID, SI: size of real spectrum, WDW: type of window function, SR: spectrum reference frequency.

Experiment	Acquisition parameters	Processing parameters
Standard ^{1}H (zg)	NS: 128; RD: 2 s; TD: 32 k; SW: 9014.42Hz/ 15.02 ppm; RG: 228	SI: 140 k; WDW: exponential; LB: 0.30 Hz; Manual phase and baseline correction; Chemical shift calibration by $CDCl_3$ (7.26 ppm)
COSY (cosygpppqf)	NS: 16; RD: 2 s; TD[F1]: 256; TD[F2]: 8 k; SW[F1]: 7201.59 Hz; SW[F2]: 7201.59 Hz	SI [F1]: 1 k; SI [F2]: 1 k; WDW: qsine; Manual phase correction; Chemical shift calibration by $CDCl_3$ (7.26 ppm)
2Q-1Q correlation	NS: 16; RD: 2 s; τ: 15 ms; t_1: 3 µs; t_2: 200 µs; G_1: 15%; G_2: -30%; TD[F1]: 128; TD[F2]: 8 k; SW[F1]: 10802.37 Hz; SW[F2]: 5405.41 Hz; O1: 2700.58 Hz	SI [F1]: 1 k; SI [F2]: 512; WDW: sine; F1 in magnitude mode; Chemical shift calibration using the SR value of 1D
3Q-1Q correlation	NS: 16; RD: 2 s; τ: 30 ms; t_1: 3 µs; t_2: 200 µs; G_1: 15%; G_2: -45%; TD[F1]: 256; TD[F2]: 8 k; SW[F1]: 16203.58 Hz; SW[F2]: 5405.41 Hz; O1: 2700.58 Hz	SI [F1]: 1 k; SI [F2]: 8 k; WDW: sine; F1 in magnitude mode; Chemical shift calibration using the SR value of 1D
4Q-1Q correlation	NS: 16; RD: 2 s; τ: 45 ms; t_1: 3 µs; t_2: 200 µs; G_1: 15%; G_2: -60%; TD[F1]: 256; TD[F2]: 8 k; SW[F1]: 16203.58 Hz; SW[F2]: 5405.41 Hz; O1: 2700.58 Hz	SI [F1]: 1 k; SI [F2]: 512; WDW: sine; F1 in magnitude mode; Chemical shift calibration using the SR value of 1D
5Q-1Q correlation	NS: 16; RD: 2 s; τ: 110 ms; t_1: 3 µs; t_2: 200 µs; G_1: 10%; G_2: -50%; TD[F1]: 256; TD[F2]: 8 k; SW[F1]: 27005.96 Hz; SW[F2]: 5405.41 Hz; O1: 2700.58 Hz	SI [F1]: 1 k; SI [F2]: 512; WDW: sine; F1 in magnitude mode; Chemical shift calibration using the SR value of 1D

^{1}H multiple quantum coherences were excited and detected with the basic pulse sequence RD-90_x-τ-180_x-τ-90_Φ-t_1-90_x-t_2-ACQ (Reddy and Caldarelli, 2010). The value of τ was optimized by eye to obtain the highest and most uniform signal intensity along the series of molecules. The phase Φ was chosen to select odd or even MQ orders (*x* or *y* for even or odd order excitation, respectively). Unwanted coherences were effectively dephased using a couple of sine-shaped pulse field gradients, G_1 and G_2, placed before and after the last 90 pulse, respectively. The ratio of the gradient pulse was selected to fulfill $G_2 = p \times G_1$ (Keeler et al., 1994), where *p* is the MQ coherence of choice (i.e. 2 for double coherence, 3 for triple and so on). In this work, the double quantum-single quantum (2Q-1Q), triple quantum-single quantum (3Q-1Q), quadruple quantum-single quantum (4Q-1Q) and quintuple quantum-single quantum (5Q-1Q)

correlations were acquired and processed with the parameters shown in Table 2.10. All NMR data processing was carried out in TOPSPIN 3.2 version (Bruker BioSpin, Reinstetten, Germany).

2.3.3. Data pre-processing and analysis

Alignment and normalization

NMR spectra were converted to a matrix of *n* rows (corresponding to *n* samples), and *m* columns (corresponding to variables). For whole blood plasma, the full resolution spectra (standard -1 to 10.5 ppm, CPMG 0.5 to 8.5 ppm and diffusion-edited -1 to 10.5 ppm) were used to construct data matrices for multivariate analysis using Amix 3.9.5 (Bruker BioSpin, Rheinstetten, Germany). The water region (4.5-5.0 ppm) was excluded as was that of ethanol (1.15-1.20 and 3.62-3.68 ppm), which was found randomly in blood samples, possibly due to skin disinfection prior to collection. Standard and CPMG spectra were aligned using recursive segment-wise peak alignment (RSPA) (Veselkov et al., 2009) to minimize chemical shift variations in some low molecular weight metabolites (e.g. valine, alanine, citrate), and data were normalized using probabilistic quotient normalization (PQN) (Dieterle et al., 2006) to account for differences in sample concentration. For plasma total lipid extracts, the full resolution standard ^{1}H spectra (0.5 to 9 ppm) were used excluding the water (1.60-1.97 ppm) and $CDCl_3$ (7.04-7.05, 7.21-7.315 and 7.46-7.472 ppm) regions. Spectra were then aligned using RSPA to minimize chemical shift variations in the phospholipids resonances and normalized using PQN, as described above for whole plasma. Alignment and normalization were performed in Matlab 7.9.0 (The MathWorks, Inc.).

Variable selection

For Studies 3 to 5 (in chapters 6 to 8), variables (spectral data points) were selected through the intersection of three parameters: a) VIP>1, b) $VIP/VIP_{cvSE}>1$ and c) $b/b_{cvSE}>1$ (Diaz et al., 2013b). Parameters b) and c) define variables with lower standard errors in relation to VIP (VIP_{cvSE}) and to b-coefficients (b_{cvSE}). The standard error obtained by cross validation (cvSE) provides a measurement of the uncertainty of a parameter based on the concept of signal to noise ratio using a jack-knife approach which estimates the statistical error by re-computing models using a random subset of samples (Andersen and Bro, 2010). This concept was applied to the VIP and b-coefficient parameters. The VIP values reflect the importance of each variable to the

projection in the low dimensional space i.e. the magnitude of the contribution of each variable to the separation, whilst b-coefficients reflects the discriminant and predictive capability of each variable i.e. their contribution for samples classification.

Multivariate analysis (MVA)

MVA was performed in NMR data of whole plasma and plasma lipid extracts after pre-processing and, in some cases, after variable selection using SIMCA-P 11.5 software (Umetrics, Umeå, Sweden). In study 3 (chapter 6), row-concatenated matrix of 2nd T plasma and urine spectra obtained from common donors (22 controls, 26 CDs, 12 T21) was considered for PLS-DA.

PCA (Jolliffe, 2002) and PLS-DA (Barker and Rayens, 2003) were performed on data scaled by different methods, for comparison purposes: unit variance (UV), Pareto and centered scaling, with UV providing the best results. Model robustness was initially evaluated in terms of Q^2 (goodness of prediction or prediction power), a parameter obtained by 7-fold cross validation in SIMCA-P 11.5 software. The corresponding loading plots of PCA and PLS-DA were back-transformed by multiplying the loading weight [w] by the variable standard deviation. PLS-DA loading plots were further colored according to each VIP, using Matlab 7.9.0. (The MathWorks, Inc.).

Additionally, Partial Least Squares (PLS) regression analysis was performed in study 4 to enable the identification of metabolites varying significantly (at 95% confidence level) between post-diagnosis GDM and controls (**X** matrix) in relation to gestational age (**y** vector).

Model validation

For PLS-DA models, Monte Carlo cross-validation (MCCV, using in-house developed software) (7 blocks, 500 runs) was carried out with recovery of Q^2 values and confusion matrices. Classification rates, specificity (spec.) and sensitivity (sens.) were computed and the predictive power of each model was further assessed using a receiver operating characteristic (ROC) map, a function of the true positive rate (TPR or sensitivity), and false positive rate (FPR or 1-specificity). PLS-DA models were considered robust when minimal overlapping of the distribution of true and permuted Q^2 was obtained (Wiklund et al., 2007; Westerhuis et al., 2008).

Univariate analysis

Relevant peaks identified in loading profiles were integrated in Amix 3.9.5 (Bruker BioSpin, Rheinstetten, Germany) and PQN normalized. For plasma lipid extracts, the $N(CH_3)_3$ choline resonance of SM (3.31 ppm) and PC (3.34 ppm) were subjected to deconvolution in order to better integrate the two close peaks. A Lorentzian curve fitting was used with a default line width of 3 Hz, in Amix 3.9.5 (Bruker BioSpin, Rheinstetten, Germany).

The individual statistical significance of each peak integral was computed by Shapiro-Wilk test (to determine if data was normally distributed), Student's t-test (used for normally distributed data) or Wilcoxon rank sum test (used for non-normally distributed data), with R-statistical software (R. D. C. Team, 2010). For each set of significant metabolites, considering a significance level of 95% ($p<0.05$), the percentage of variation and the effect size value were calculated. The percentage of variation and the uncertainty were calculated according to Equation 2.1 and Equation 2.2, respectively, where $\bar{X}$ denotes the average value, S the standard deviation and n the sample numbers for both groups a and b. The effect size (E.S.) and the respective standard error were calculated according to Equation 2.3 and Equation 2.4, respectively, where $\bar{X}$ denotes the average value, S the standard deviation and n the sample numbers for both groups a and b (Berben et al., 2012). The distributions of peak integrals were also accessed with aid of box-and-whisker plots (boxplots) for group comparison.

$$\%\, variation = \frac{(\overline{X}_b - \overline{X}_a)}{\overline{X}_b} \times 100 \qquad \text{Equation 2.1}$$

$$\%\, uncertainty = \frac{\sqrt{\frac{{S_a}^2}{n_a} + \frac{{S_b}^2}{n_b}}}{\frac{(n_a \times \overline{X}_a + n_b \times \overline{X}_b)}{(n_a + n_b)}} \qquad \text{Equation 2.2}$$

$$Effect\ size\ (E.S.) = \frac{(\overline{X}_b - \overline{X}_a)}{\sqrt{(n_a - 1){S_a}^2 + (n_b - 1){S_b}^2/(n_a + n_b - 2)}} \qquad \text{Equation 2.3}$$

$$E.S._{standard\ \ error} = 1.96 \times \sqrt{\frac{n_a + n_b}{n_a n_b} + \frac{ES^2}{2(n_a + n_b)}} \qquad \text{Equation 2.4}$$

For study 1, the coefficient of variation (CV) was computed for each variable (PQN normalized data point) as the ratio between the standard deviation and the average

value; in this calculation, the noise was estimated and removed from the data according to the exclusion criteria if $({y_i}^{max} - {y_i}^{min}) < 3P_{10}(\sigma)$, where ${y_i}^{max}$ and ${y_i}^{min}$ are respectively the maximum and minimum intensities for each data point and $P_{10}(\sigma)$, is the 10th percentile of the standard deviation for all data points, based on references (Golotvin and Williams, 2000; Parsons et al., 2009).

For study 1 and 2, the metabolite variations were also expressed in a heatmap representation with integral values normalized to unity and shown in a qualitative color scale, in R-statistical software. In study 3, 4 and 5, each integrals' effect size and statistical significance (p-value) were further represented by means of a Volcano plot which is a graphical summary of the effect size vs. -log(p-value) (Cui and Churchill, 2003). Additionally, in study 4, Bonferroni correction was used to adjust p-values for multiple comparisons by setting the significance cut-off to α/n with α=0.05 and n=number of metabolites/resonances simultaneously tested.

Correlation studies

Statistical total correlation spectroscopy (STOCSY) was performed for each unassigned NMR resonance and for the CH_3 resonances (at 0.84, 0.86 and 0.87 ppm) of lipoprotein particles in order to identify peaks arising from the same molecule and/or metabolically related compounds. The correlation coefficient and the covariance were calculated between each resonance and the complete spectrum matrix. The covariance was then plotted and colored according to the correlation coefficient (r), in Matlab 7.9.0 (The MathWorks, Inc.) (Cloarec et al., 2005). In order to facilitate the identification of the correlated peaks, the 1D STOCSY was plotted without any threshold.

For study 2 (chapter 5), in order to define a more complete picture of pregnancy biofluid metabolome, plasma changes were correlated, using Pearson Product-Moment correlation analysis, to those observed in urine (Diaz et al., 2013a) collected for a partially overlapped cohort (in a total of 69 subjects, 30 provided both plasma and urine: 15 in the NP group and the remaining distributed by the different pregnancy trimesters) recruited with the same criteria as for plasma. Intra-biofluid correlations (plasma/plasma and urine/urine) were also calculated. A correlation coefficient (r) and significance threshold of $|r| \geq 0.6$ and $p < 0.05$ were applied (to select the top 3% of correlation values) and correlations were confirmed by scatter plotting each pair of correlated integrals in Matlab 7.9.0 (The MathWorks, Inc.). These pairs were then used to compute a correlation network, using the Gephi 0.8.2 software (Bastian et al., 2009).

In study 3 (chapter 6), the plasma and urine metabolites found to change in CD cases compared to controls (26 CD and 22 controls from common donors) were used for intra- and inter-biofluid Pearson Product-Moment correlation analysis (R-statistical software). Correlation coefficients (r) ranged from -0.609 to 0.988 and correlations were confirmed by scatter plot inspection.

2.4. Profiling of maternal plasma by Hydrophilic Interaction Liquid Chromatography-Mass Spectrometry (HILIC-LC/MS) phospholipid analysis

2.4.1. Sample preparation

Blood plasma pools were thawed at room temperature and prior to phospholipids extraction, 5 µg (7.4 pM) of an internal standard (1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC)) was added to 100 µL of each pooled sample. Phospholipids extraction was performed in a HybridSPE-Phospholipid (HybridSPE-PL) column (Sigma Aldrich, Portugal), as illustrated in Figure 2.4.

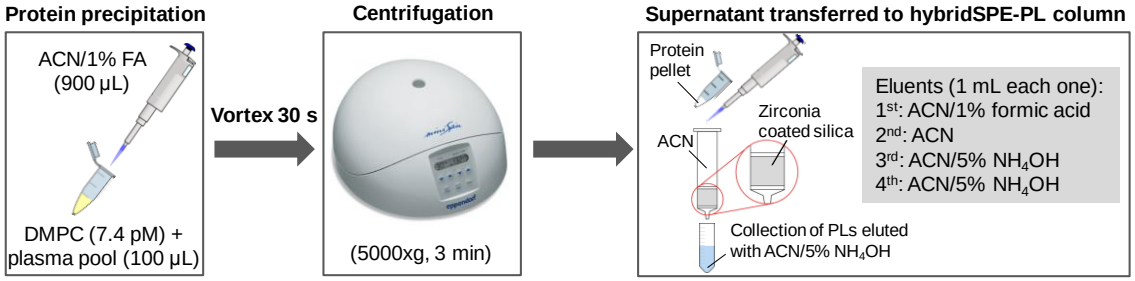


Figure 2.4. Schematic representation of extraction of phospholipids from pooled plasma samples using the hybridSPE-PL column.

Briefly, 900 µL of acetonitrile (ACN) with 1% formic acid were added to 100 µL of pooled sample, vortexed (30 s) and centrifuged (5000 *g*, 3 min). The supernatant was transferred to the HybridSPE-PL column, followed by addition of ACN with 1% formic acid (1 mL) and ACN (1 mL). The retained phospholipids were eluted with two consecutive aliquots of ACN with 5% ammonium hydroxide (NH_4OH) (1 mL). The flow-through was collected, dried, and stored at -20ºC until analysis. The interaction between the zirconia coated silica in the HybridSPE-PL column and plasma phospholipids is based on Lewis acid-base chemistry and the recovery of phospholipids is over 95% as has been thoroughly discussed in reference (Lu and Ye,

2010). Before HILIC-LC/MS analysis, phospholipid extracts were thawed and resuspended in 200 µL of $CHCl_3$, from which 2 µL were taken and dried under nitrogen. After drying, the aliquot of phospholipid extract was diluted with 90 µL of mobile phase B (60% acetonitrile, 40% methanol with 1 mM ammonium acetate).

2.4.2. HILIC-LC/MS analysis

Phospholipid classes were separated by HILIC-LC/MS, performed on an HPLC system (Waters Alliance 2690) coupled to an electrospray (ESI) linear ion trap mass spectrometer LXQ (ThermoFinnigan, San Jose, CA, USA). The mobile phase A consisted of 10% water, 55% acetonitrile and 35% (v/v) methanol with 1 mM ammonium acetate. The mobile phase B consisted of 60% acetonitrile, 40% methanol with 1 mM ammonium acetate. Phospholipid extracts diluted in the mobile phase B (15 µL) were introduced into an Ascentis Si HPLC Pore column (15 cm × 1.0 mm, 3 µm, Sigma-Aldrich). A pre-column split (Acurate, LC Packings, USA) was used in order to obtains a flow rate of 50 µL min^{-1}. The solvent gradient was programmed as follows: 100% of B was held isocratically for 15 min, followed by a linear increase to 100% A during 5 min, and held isocratically for 25 min, returning to the initial conditions in 5 min.

Phospholipid detection and identification was carried out under positive ionization mode on an ESI linear ion trap mass spectrometer (ThermoFinnigan, San Jose, CA, USA). ESI conditions on linear ion trap mass spectrometer were as follows: electrospray voltage was 5 kV; capillary temperature was 275ºC and the sheath gas flow was 8 units. To obtain the product-ion spectra of the major components during LC experiments, cycles consisting of one full scan mass spectrum and four data-dependent MS/MS scans were repeated continuously throughout the experiments with the following dynamic exclusion settings: repeat count 1; repeat duration 30 s; exclusion duration 45 s. MS data were collected in centroid mode in the m/z range 80-2000. Data acquisition was carried out with an Xcalibur data system (V2.0). Phospholipid classes were identified using the MS/MS data and the LIPID MAPS MS predict tool (Fahy et al., 2007).

2.4.3. Data pre-processing and analysis

Data filtering, peak detection, chromatogram deconvolution and chromatogram alignment (i.e. matching of equivalent peaks across multiple samples) were performed

using MZmine (version 2.10, MZmine Development Team) (Pluskal et al., 2010). The parameters were set as follow: RT range 7.0-50.0 min, m/z range 450.0 – 950.0, MS data noise level 1.5E2; m/z tolerance 0.5 m/z or 300 ppm; minimum time span of chromatographic peaks 1.0 min; and a minimum intensity of a peak in chromatogram of 1.5E2. The area of each peak, after being recognized and aligned, was normalized to the area of the internal standard DMPC added to each plasma pool before phospholipid extraction. Isotopic peaks were excluded for processing.

The resulting data, m/z-RT pair, sample name and normalized ion area were subjected to multivariate analysis. PCA and PLS-DA were performed on data scaled by unit variance (UV), using SIMCA-P 11.5 software (Umetrics, Umeå, Sweden). For PLS-DA models, MCCV (7 blocks, 500 runs, using in-house-developed software) was carried out with recovery of Q^2 values and computation of classification rates, sensitivity and specificity, as described above for NMR data. Relevant m/z-RT pairs identified in loadings profiles with VIP higher than 1, were analysed using univariate analysis tests: Shapiro-Wilk normality test, Student's t-test or Wilcoxon test, with R-statistical software. Effect size (E.S.) was computed for m/z pairs with statistically significant differences between control and diseases ($p<0.05$), corrected for small sample numbers (Equation 2.5 and Equation 2.6).

$$E.S._{corrected} = J \times ES, \qquad where\ J = 1 - \frac{3}{4(n_a + n_b - 2) - 1} \qquad \text{Equation 2.5}$$

$$E.S._{standard\ \ error\ \ corrected} = J \times \sqrt{\frac{n_a + n_b}{n_a n_b} + \frac{ES^2}{2(n_a + n_b)}} \qquad \text{Equation 2.6}$$

The correction is performed through the multiplication of E.S. and $E.S._{standard\ error}$ (Equation 2.3 and Equation 2.4, defined above for NMR data) by a factor J defining two new equations, where n is the sample numbers for both groups a and b (Berben et al., 2012).

Chapter 3. NMR plasma metabolome and lipidome

This chapter begins by presenting the NMR plasma metabolome obtained by 1D and 2D experiments of whole blood plasma and the resulting complete list of metabolites identified. Secondly, the NMR plasma lipidome characterization through 1D and 2D NMR experiments is presented, comprising results of the application of MQ spectroscopy for lipid assignment in complex samples.

3.1. NMR spectra of blood plasma and spectral assignment

Blood plasma is the most important transport medium in the body containing metabolites that are intermediates and end products of metabolism in cells, tissues, organs and other biofluids. In order to characterize blood plasma composition, plasma samples were analysed at 500 MHz and prepared as described in Chapter 2 (section 2.3). The presence of low and high M_w compounds in plasma enforces the use of standard 1D, CPMG (or T_2-edited) and diffusion-edited spectra to characterize the NMR plasma metabolome. Figure 3.1 shows an example of the typical ^{1}H NMR spectra obtained for a healthy non-pregnant woman. The standard 1D spectrum (Figure 3.1a) is dominated by the broad resonances of large molecules, mostly lipoproteins and proteins, as well as by many signals from low M_w metabolites. A high number of signals arise from the different types of protons in the fatty acids (FA), free cholesterol (FC), esterified cholesterol (EC), triglycerides (TG) and phospholipids (PL) present in lipoprotein particles. In addition, broad signals from the N-acetyl groups of glycoproteins, lysyl groups of albumin (most abundant plasma protein) and NH from total proteins, are detected. Furthermore, several resonances from small metabolites can be identified in the standard 1D spectrum, such as glucose, some amino acids (valine, alanine, arginine, creatine, glutamine, threonine), organic acids (lactate, acetate, pyruvate, citrate), ketone bodies (acetone, acetoacetate), creatinine, betaine and trimethylamine N-oxide (TMAO). These metabolites are more clearly detected in the CPMG spectrum (Figure 3.1b), due to the substantial attenuation of signals from proteins and lipoproteins, with the addition of others present in lower concentrations (leucine, isoleucine, glycine, lysine, histidine, tyrosine, formate, dimethyl sulfone). The diffusion-edited spectrum (Figure 3.1c) is especially useful to study the lipidic components of plasma without the interference of narrow peaks from low M_w metabolites (e.g., glucose, lactate). This spectrum is dominated by the CH_3, $(CH_2)_n$ and unsaturated protons from fatty acids, $N(CH_3)_3$ choline resonances of choline-containing PL, and glyceryl protons in PL and TG. The N-acetyl glycoprotein and albumin-lysyl

signals are also observed, as well as the large signal of NH from total proteins at δ 5.5-11.0.

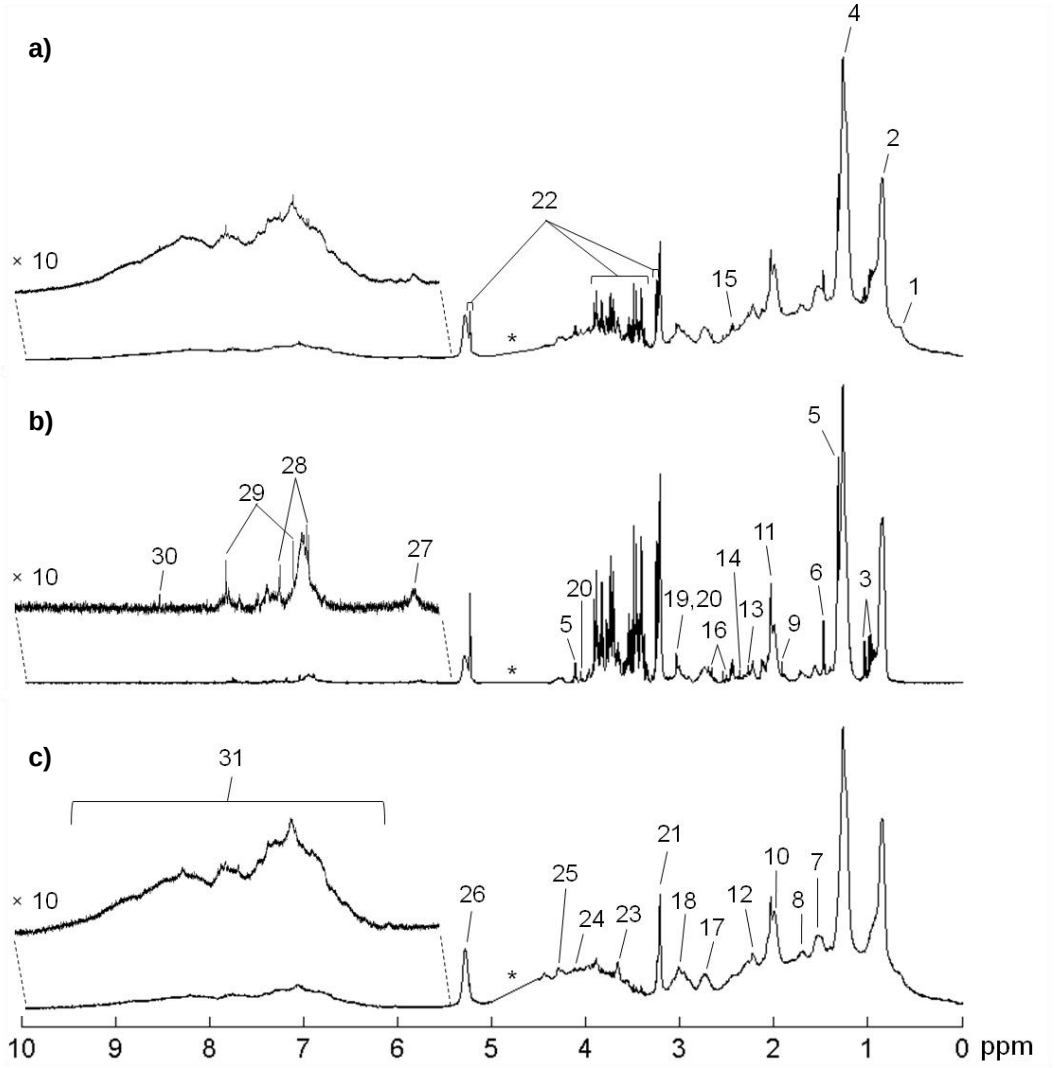


Figure 3.1. 1D ^{1}H NMR 500 MHz spectra of plasma of a healthy non-pregnant woman: a) standard, b) CPMG (T_2-edited), and c) diffusion-edited spectra. Assignment: 1- $C18H_3$ cholesterol, 2- CH_3 lipids, 3- valine, 4- $(CH_2)_n$ lipids, 5- lactate, 6- alanine, 7- $C\underline{H}_2CH_2CO$ lipids, 8- $C\underline{H}_2CH_2CH$=, 9- acetate, 10- $C\underline{H}_2CH$= lipids, 11- N-acetyl glycoproteins, 12- CH_2CO lipids, 13- acetoacetate, 14- pyruvate, 15- glutamine, 16- citrate, 17- =CHC$\underline{H}_2$CH= lipids, 18- albumin-lysyl, 19- creatine, 20- creatinine, 21- $N(CH_3)_3$ choline of PL, 22- α- and β- glucose, 23- $C\underline{H}_2N(CH_3)_3$ choline of PL, 24- glyceryl-C1,3H, 25- glyceryl-C1,3H', 26- HC=CH lipids, 27- urea, 28- tyrosine, 29- histidine, 30- formate, 31- NH proteins. *: cut off spectral region corresponding to water signal.

In order to improve the assignment of the ^{1}H NMR spectra of plasma, 2D ^{1}H-^{1}H TOCSY, ^{1}H-^{13}C HSQC and J-resolved experiments were acquired. The ^{1}H-^{1}H TOCSY spectrum of blood plasma for the same healthy non-pregnant woman is shown in Figure 3.2a. This spectrum is dominated by broad signals arising from the high M_w metabolites described above, but still enabling the confirmation of spin systems from

metabolites (e.g., leucine, lysine, β-hydoxybutyrate) present in low concentrations and obscured in the 1D spectra. The ^{1}H-^{13}C HSQC, in Figure 3.2b, contains only the signals from lipidic resonances, N-acetyl glycoproteins, albumin-lysyl and glucose. Finally, the *J*-res experiment (not shown) was also acquired, enabling the further confirmation of peak assignments from low M_w compounds through their multiplicity and *J*-coupling constants. Furthermore, peak assignment was confirmed by comparison with several databases, such as the BMRB (Seavey et al., 1991), HMDB (Wishart et al., 2007) and Bruker Biorefcode spectral database (courtesy of Bruker Biospin, Rheinsteten, Germany), and previous reports (Nicholson et al., 1995; Liu et al., 2002; Engelke et al., 2005).

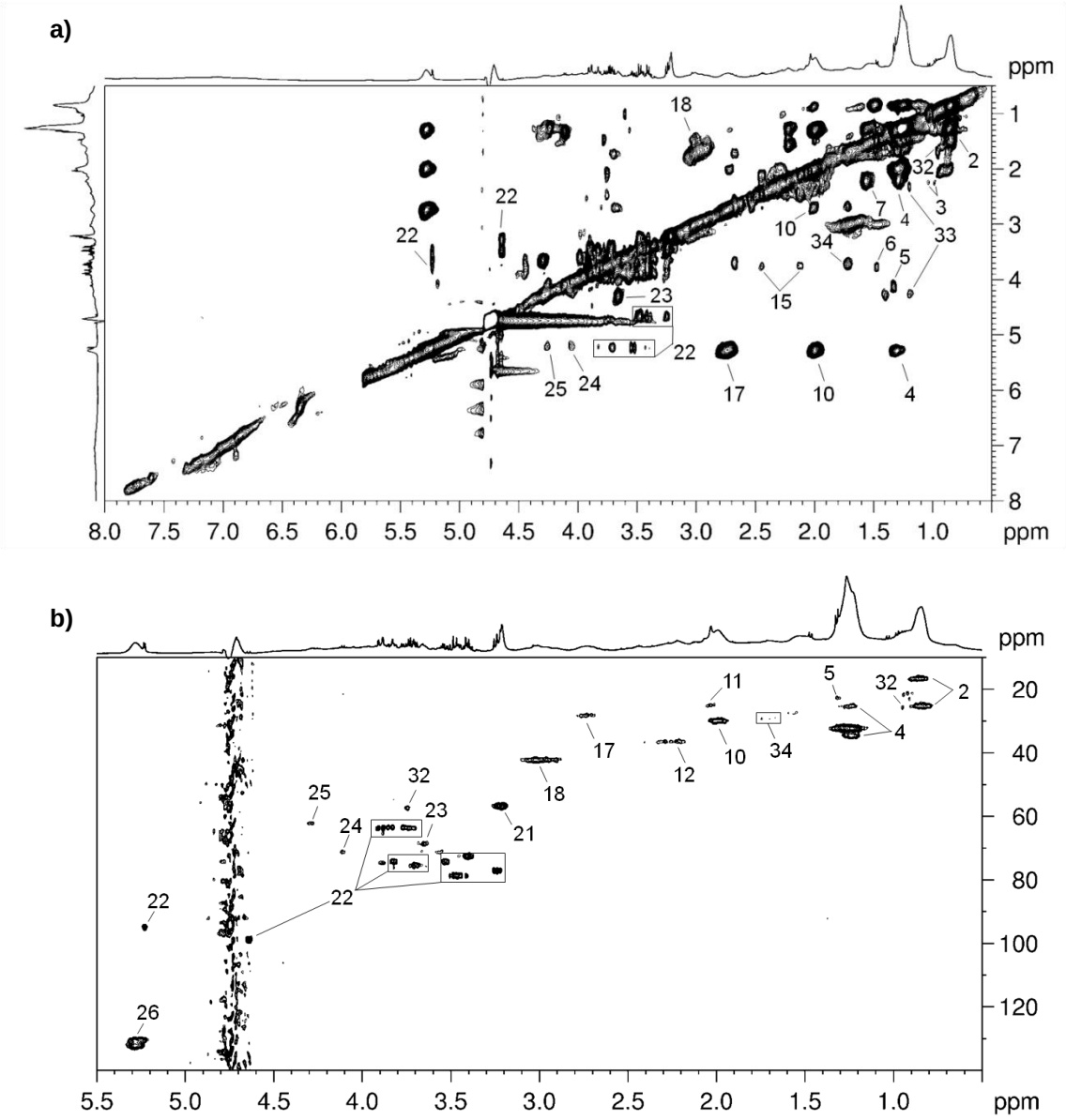


Figure 3.2. 2D 500 MHz spectra of plasma from a healthy non-pregnant woman: a) TOCSY, and b) ^{1}H-^{13}C HSQC. Peak assignments correspond to that shown in Figure 3.1, with the addition of 32- leucine, 33- β-hydroxybutyrate, 34- lysine.

Overall, 31 low M_w metabolites, lipids and proteins (albumin and N-acetyl glycoproteins) were identified in plasma, all having been previously identified in human blood plasma/serum by NMR (Nicholson et al., 1995; Liu et al., 2002; Engelke et al., 2005), except 1,5-anhydroglucitol (present at 73.0-243.0 μM in blood (Wishart et al., 2007)) reported only in a GC-MS study (Scholtens et al., 2014). A comprehensive list of ^{1}H and ^{13}C (measured in HSQC) NMR assignment of plasma metabolites is presented in Table 3.1, along with the corresponding chemical shifts and multiplicity.

Table 3.1. 500 MHz ^{1}H and ^{13}C NMR assignment of metabolites present in human plasma samples (average pH 7.86). s: singlet, d: duplet, dd: doublet of doublets, dt: doublet of triplets, t: triplet, q: quartet, m: multiplet, br: broad. [a] identified here for the first time to our knowledge, by NMR spectroscopy.

Compound	δ ^{1}H in ppm (multiplicity, assignment) / δ ^{13}C in ppm
Acetate	1.91 (s,CH_3)
Acetoacetate	2.27 (s,CH_3)
Acetone	2.22 (s, CH_3)
Alanine	1.46 (d,CH_3); 3.75 (q, α-CH)
Albumin lysyl	2.90 (t, ϵCH_2)/42.14; 2.96 (t, εCH_2)/42.14; 3.01 (t, εCH_2)/42.14
1,5-anhydroglucitol [a]	3.26 (t, C1H); 3.34 (m, C4,5H); 3.43 (m, C3H); 3.58 (m, C2H); 3.68 (dd, C6H); 3.87 (d, C6H'); 3.97 (dd, C1H')
Betaine	3.29 (s, NCH_3)
Cholesterol	0.65 (m, C18H)
Citrate	2.51 (d, α,βCH); 2.66 (d, α',β'CH)
Citrulline	1.57 (m, γCH_2); 1.86 (m, βCH_2); 3.15 (t, γ'CH_2); 3.70 (m, αCH)
Creatine	3.03 (s, NCH_3); 3.92 (s, NCH_2)
Creatinine	3.04 (s, NCH_3); 4.05 (s, NCH_2)
Dimethylsulfone	3.14 (s, CH_3)
Formate	8.45 (s, CH)
α-glucose	3.41 (t, C5H)/72.40; 3.53 (dd, C2H)/74.30; 3.71 (t, C3H)/75.57; 3.76 (dd, C6H)/63.48; 3.83 (m, C4H)/74.18; 3.83 (m, C6H')/63.40; 5.23 (d, C1H)/94.96
β-glucose	3.23 (dd, C2H)/77.10; 3.40 (t, C5H)/72.40; 3.47 (dd, C4H)/78.70; 3.49 (t, C3H)/78.49; 3.71 (dd,C6H')/63.70; 3.89 (m, C6H)/63.70; 4.64 (d, C1H)/98.74
Glutamine	2.12 (m, βCH_2); 2.44 (m, γCH_2); 3.75 (t, αCH)
Glycerol	3.55 (dd, C1,3H); 3.65 (dd, C1,3H'), 3.78 (m, C2H)
Glycine	3.55 (s, αCH)
β-hydroxybutyrate	1.19 (d, γCH_3), 2.30 (m, CH_2), 2.40 (m, CH_2), 4.15 (m, CH)
Histidine	3.10 (dd, βCH_2); 3.19 (dd, β'CH_2); 3.91 (dd, αCH); 7.04 (s, C4H, ring); 7.74 (s, C2H, ring)
Isoleucine	0.93 (t, γCH_3); 1.00 (d, βCH_3)
Lactate	1.32 (d, CH_3); 4.10 (q, CH)/22.61

Table 3.1. (cont.)

Compound	δ ^{1}H in ppm (multiplicity, assignment) / δ ^{13}C in ppm
Leucine	0.95 (d, δ,δ'CH)/25.65; 1.70 (m, β,γCH); 3.72 (t, αCH)/57.42
Lipids	0.85 (br, CH_3)/16.58, 25.20; 1.25 (br, $(CH_2)_n$)/25.19, 32.25, 34.55; 1.53 (br, C$\underline{H}_2$$CH_2CO$); 1.70 (br, C$\underline{H}_2$$CH_2C$=C); 1.99 (br, CH_2C=C)/29.79; 2.22 (br, CH_2CO)/36.43; 2.74 (br, C=CCH_2C=C)/28.13; 3.21 (br, N$(CH_3)_3$ choline)/56.58; 3.66 (br, C$\underline{H}_2$-N$(CH_3)_3$ choline)/68.70; 4.11 (br, glyceryl-C1,3H) /71.27, 4.24 (br, glyceryl-C1,3H')/62.14, 5.19 (br, glyceryl-C2H), 5.29 (m, HC=CH)/131.20
Lysine	1.71 (m, δCH_2); 1.89 (m, βCH_2); 2.99 (t, εCH_2);
Methanol	3.36 (s, CH_3)
N-acetyl glycoproteins	2.03 (s,CH_3)/24.93
Phenylalanine	3.12 (dd, βCH_2); 3.26 (dd, β'CH_2); 3.97 (dd, αCH); 7.31 (m, C2H/C6H); 7.36 (m, C4H); 7.41 (m, C3H/C5H)
Proline	2.01 (m, γCH_2); 2.07 (m, βCH_2); 2.34 (m, β'CH_2); 3.34 (dt, δCH_2); 3.45 (m, δ'CH_2); 4.12 (dd, αCH)
Pyruvate	2.36 (s, CH_3)
Threonine	1.32 (d, γCH); 3.57 (d, αCH); 4.24 (m, βCH)
Trimethylamine N-oxide (TMAO)	3.27 (s, NCH_3)
Tyrosine	3.06 (dd, βCH_2); 3.16 (dd, β'CH_2); 3.94 (dd, αCH); 6.89 (d, C3H/C5H); 7.18 (d, C2H/C6H)
Valine	0.98 (d, γCH_3); 1.03 (d, γ'CH_3); 2.26 (m, βCH); 3.60 (d, αCH)
Urea	5.77 (br, NH_2)

Regarding the identification of different lipoprotein particles by NMR, some reports have identified some degree of peak resolution in the CH_3 and $(CH_2)_n$ resonances, without physical isolation, due the slightly chemical shift differences in HDL, LDL and VLDL (Otvos et al., 1991; Ala-korpela et al., 1994; Liu et al., 2002). In this work, statistical total correlation spectroscopy (STOCSY) analysis was found to be a useful tool to identify the ^{1}H resonances arising from HDL and LDL+VLDL as will be shown in Chapter 4 (section 4.2) and 5 (section 5.1).

3.2. NMR spectra of plasma lipid extracts and spectral assignment

3.2.1. 1D and 2D NMR

In order to better characterize the lipidic components of blood plasma, an extraction procedure was applied followed by 500 MHz 1D and 2D NMR acquisition, as described in Chapter 2 (section 2.3). The lipid extraction procedure enables the solubilisation of all lipid classes present in lipoprotein particles. Different extraction procedures for plasma lipid extraction were previously analysed and compared by NMR in a master thesis (Almeida, 2014), unveiling that the most suitable method is the MTBE-methanol (easier extraction and more reproducible). A recent LC-MS study (Sarafian et al., 2014) also reported broad lipid coverage with recoveries of about 60%, and 99% protein removal using this method for plasma lipid extraction. Figure 3.3 shows the comparison of the diffusion-edited spectrum of plasma and the corresponding ^{1}H spectrum of lipid extract, for a healthy pregnant woman in 1st T. Compared with whole plasma, the lipid extract shows better resolution of lipid classes but still with high overlap, namely in the fatty acyl and cholesterol (free and esterified) resonances. The region δ 0.65-1.40 is dominated by the ^{1}H resonances of free cholesterol (FC) and esterified cholesterol (EC), partially superimposed in the CH_3 resonance of FA, as well as the $(CH_2)_n$ signal of FA. Distinct resonances from FC and EC can be found at δ 1.01 and δ 3.53, and δ 1.02 and δ 4.61, respectively. The two most abundant PL classes in plasma, PC (1974 μM) and SM (303 μM) (Quehenberger et al., 2010), are clearly detected in the region δ 3.20-3.40, while in the region δ 3.4-4.7, it is possible the identification of glyceryl resonances of total PL and TG. Furthermore, the assignment of several unsaturated FA resonances was possible through 2D spectra (TOCSY and HSQC, Figure 3.4a and b) and 1D STOCSY (Figure 3.4c), namely oleic acid (18:1), linoleic acid (18:2), arachidonic acid (20:4) and docosahexaenoic acid (22:6). These unsaturated FA are the most abundant in plasma, with reported percentages of contribution to the total plasma FA composition of 24.7% for 18:1, 21.2% for 18:2, 5.1% for 20:4 and 1.3% for 22:6 (Glaser et al., 2010). In Figure 3.4c, the resonance of 20:4 FA at δ 2.82 was correlated with the whole spectrum and peaks arising from the same molecule are clearly observed with high covariance and correlation (red and orange). STOCSY added information mainly in the identification of CH_3 resonances of polyunsaturated FA, which were highly overlapped even in the 2D spectra. In some cases, lipid assignments were confirmed by ^{1}H NMR experiments of pure standards (18:2 CE, 20:4 CE, 16:0/18:2 PC, d18:1/16:0 SM) and computer simulations (^{1}H and ^{13}C NMR

prediction tool of ChemBioDraw Ultra 14.0, CambridgeSoft Corporation, PerkinElmer, Inc.).

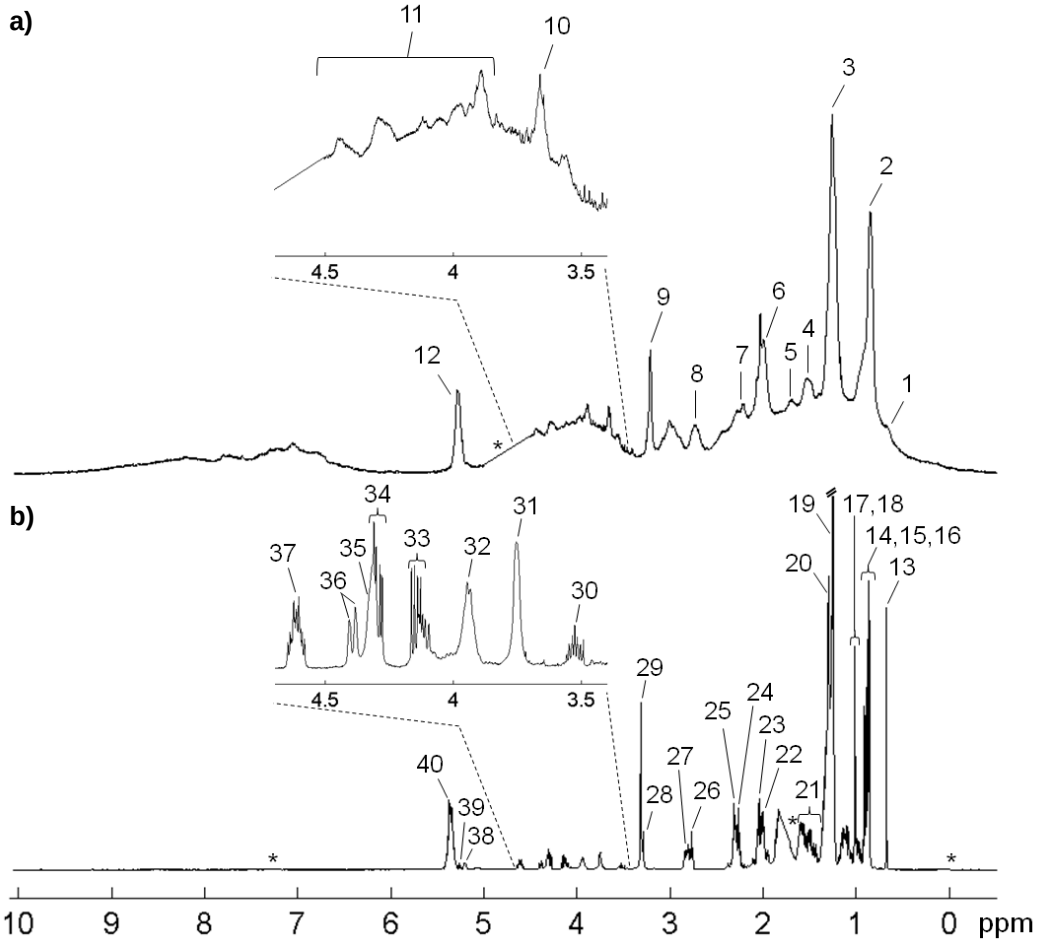


Figure 3.3. a) 500 MHz diffusion-edited spectrum of blood plasma and b) 500 MHz ^{1}H spectrum of plasma lipid extract of a healthy pregnant woman in the 1st trimester. Assignment: 1- C18H3 cholesterol, 2- CH_3 lipids, 3- $(CH_2)_n$ lipids, 4- $C\underline{H}_2CH_2CO$ lipids, 5- $C\underline{H}_2CH_2CH=$ lipids, 6- $C\underline{H}_2CH=$ lipids, 7- CH_2CO lipids, 8- $=CHC\underline{H}_2CH=$ lipids, 9- $N(CH_3)_3$ choline of PL, 10- $C\underline{H}_2N(CH_3)_3$ choline of PL, 11- glyceryl protons of lipids, 12- HC=CH lipids, 13- $C18H_3$ FC and EC, 14- $C26H_3$ and $C27H_3$ FC and EC, 15- CH_3 FA, 16- $C21H_3$ FC and EC, 17- $C19H_3$ FC, 18- $C19H_3$ EC, 19- $(CH_2)_n$ FA, 20- $=CHCH_2(C\underline{H}_2)_n$ FA, 21- multiple protons from FC and EC, 22- $C\underline{H}_2CH_2CH=$ 18:1 FA, 23- $C\underline{H}_2CH=$ FA, 24- C4H' FC, 25- CH_2CO FA, 26- $=CHC\underline{H}_2CH=$ 18:2 FA, 27- $=CHC\underline{H}_2CH=$ 20:4 FA, 28- $N(CH_3)_3$ choline SM, 29- $N(CH_3)_3$ choline PC, 30- C3H FC, 31- $CH_2N(CH_3)_3$ choline PL, 32- glyceryl-$C3H_2$ PL, 33- glyceryl-C1,3H TG, 34- glyceryl-C1,3H' TG, 35- $POCH_2$ PL, 36- glyceryl-$C1H_2$ PL, 37- C3H EC, 38- glyceryl-C2H PL, 39- glyceryl-C2H TG, 40- HC=CH FA. *: cut off spectral region corresponding to water signal in a), and water and $CDCl_3$ in b).

Overall, four distinct unsaturated FA, total saturated FA, FC, EC, total PL, PC, SM and TG were identified in plasma lipid extracts. All these lipidic components were detected before in plasma and serum lipid extracts by NMR (Willker and Leibfritz, 1998; Oostendorp et al., 2006; Tukiainen et al., 2008; Srivastava et al., 2010). Table 3.2 lists the ^{1}H and ^{13}C (from HSQC) resonances identified for each lipid class and individual unsaturated FA, along with the corresponding multiplicity and assignment.

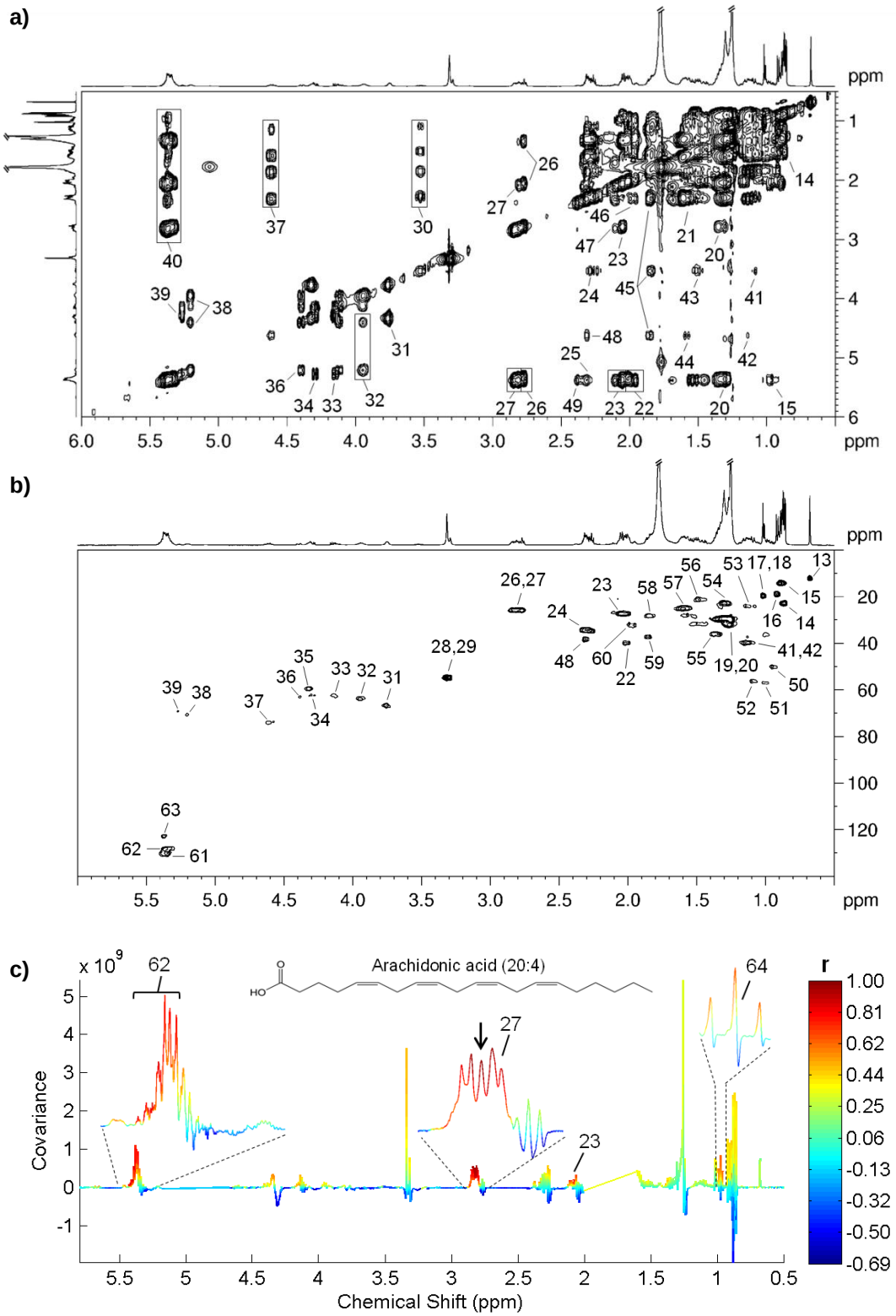


Figure 3.4. 500 MHz a) TOCSY, b) ^{1}H-^{13}C HSQC and c) STOCSY spectra of lipid extract of a healthy pregnant woman in the 1st T. Peak assignments correspond to that shown in Figure 3.3, with the addition of 41- C1H FC, 42- C2H EC, 43- C2H FC, 44- C2H EC, 45- C25H, C15H' and C16H' FC and EC, 46- C7H EC, 47- C$\underline{H}_2$CH= 20:4 FA, 48- C4H' EC, 49- CH$_2$CO 22:6 FA, 50- C9H FC and EC, 51- C14H FC and EC, 52- C17H FC and EC, 53- C23H FC and EC, 54- =CHCH$_2$(C$\underline{H}_2$)$_n$ 18:1, 55- =CHCH$_2$(C$\underline{H}_2$)$_n$ 20:4 and 18:2, 56- C2H FC, 57- C2H EC, 58- C4H FC, 59- C4H EC, 60- C7H FC and EC, 61- HC=CH 18:1 and 18:2 FA, 62- HC=CH 20:4, 63- HC=CH 22:6, 64- CH$_3$ 20:4 FA.

Table 3.2. 500 MHz ^{1}H and ^{13}C NMR assignment of lipids present in human plasma lipid extracts. s: singlet, d: duplet, dd: doublet of doublets, t: triplet, q: quartet, m: multiplet, br: broad.

Compound	**δ ^{1}H in ppm (multiplicity, assignment) / δ ^{13}C in ppm**
Arachidonic acid (20:4)	0.98 (t, CH_3); 1.34 (m, $=CHCH_2(C\underline{H}_2)_n$)/35.95; 2.07 (q, $C\underline{H}_2CH=$)/27.19; 2.81 (m, $=CHC\underline{H}_2CH=$)/25.63; 5.37 (m, HC=CH)/128.28
Docosahexaenoic acid (22:6)	2.38 (t, CH_2CO)/33.20; 2.82 (m, $=CHC\underline{H}_2CH=$); 5.36 (m, HC=CH)/122.66
Esterified cholesterol	0.68 (s, $C18H_3$)/11.84; 0.86 (d, $C26H_3$)/22.66; 0.87 (d, $C27H_3$)/22.66; 0.91 (d, $C21H_3$)/18.70; 0.95 (m, C9H)/49.96; 0.99 (m, C14H)/56.91; 1.02 (s, $C19H_3$)/19.38; 1.10 (m, C17H)/56.05; 1.12 (m, C23H)/23.90; 1.12 (m, C24H)/39.70; 1.13 (m, C22H)/37.17; 1.15 (m, C1H)/39.65; 1.35 (m, C20H)/35.94; 1.58 (m, C2H)/24.89; 1.60 (m, C15H, C16H)/24.83; 1.84 (t, C4H)/28.03; 1.87 (t, C15H', C16H'); 1.95 (t, C7H)/32.05; 1.99 (t, C7H')/31.86; 2.31 (t, C4H')/34.41; 4.61 (m, C3H)/73.86
Free cholesterol	0.68 (s, $C18H_3$)/11.84; 0.86 (d, $C26H_3$)/22.66; 0.87 (d, $C27H_3$)/22.66; 0.91 (d, $C21H_3$)/18.70; 0.95 (m, C9H)/49.96; 0.99 (m, C14H)/56.91; 1.01 (s, $C19H_3$)/19.38; 1.07 (m, C1H)/39.65; 1.10 (m, C17H)/56.05; 1.12 (m, C23H)/23.90; 1.12 (m, C24H)/39.70; 1.13 (m, C22H)/37.17; 1.35 (m, C20H)/35.94; 1.49 (m, C11H)/21.04; 1.51 (m, C2H)/21.06; 1.60 (t, C16H)/24.83; 1.84 (m, C4H)/28.08; 1.87 (t, C16H);1.95 (t, C7H)/32.05; 1.99 (t, C7H')/31.86; 2.26 (t, C4H')/34.41; 3.53 (m, C3H)
Linoleic acid (18:2)	0.98 (t, CH_3); 1.33 (m, $=CHCH_2(C\underline{H}_2)_n$)/35.95; 2.04 (q, $C\underline{H}_2CH=$)/27.19; 2.77 (t, $=CHC\underline{H}_2CH=$)/25.71; 5.34 (m, HC=CH)/130.00
Oleic acid (18:1)	0.95 (t, CH_3); 1.30 (m, $(CH_2)_n$)/22.60; 2.00 (m, $C\underline{H}_2CH_2CH=$)/39.70; 5.34 (m, HC=CH)/130.00
Phospholipids	3.29 (s, $N(CH_3)_3$ SM)/54.43; 3.31 (s, $N(CH_3)_3$ PC)/54.54; 3.75 (br, $C\underline{H}_2N(CH_3)_3$)/66.47; 3.94 (br, glyceryl-$C3H_2$)/59.43; 4.32 (br, PO-CH_2)/59.44; 4.38 (br, glyceryl-$C1H_2$)/62.90; 5.20 (br, glyceryl-C2H)/70.46
Saturated fatty acids	0.89 (br, CH_3)/14.12; 1.25 (br, $(CH_2)_n$)/29.6; 1.26(br, $(CH_2)_n$)/31.80; 1.60 (m, $C\underline{H}_2CH_2CO$)/24.88; 2.30 (m, CH_2CO)/34.14, 38.22;
Triglycerides	4.14 (dd, glyceryl-C1,3H)/62.28; 4.29 (dd, glyceryl-C1,3H')/62.14; 5.26 (m, glyceryl-C2H)/62.04

3.2.2 Multiple quantum (MQ) spectroscopy

In an attempt to overcome the complex overlap of ^{1}H signals in complex mixtures, such as plasma lipid extracts, the potential of multiple quantum (MQ) spectroscopy to generate a specific fingerprint of a given molecular structure was evaluated. This work was performed as courtesy of Institut des Sciences Moléculaires de Marseille (UMR 7313 iSm2), Aix Marseille Université, and the detailed experimental procedure can be found in Chapter 2 (section 2.3). A set of pure lipid standards were chosen based on the lipidic composition of plasma (more abundant ones), comprising a total of 6 lipid standards, 4 PL with the same fatty acyl chains (16:0 LPC, 16:0/18:2 PC, 16:0/18:2 PE,

d18:1/16:0 SM), and 2 EC (18:2 EC and 20:4 EC). Due the high similarity in the molecular structures, it was expected that the proton resonances in 1D and 2D spectra would be largely overlapped as shown above for plasma lipid extracts. An example of a ^{1}H-^{1}H COSY spectrum of 18:2 EC is shown in Figure 3.5a. This experiment was used for assignment of each pure molecule giving only the information about the coupled spins, thus simpler than TOCSY (all spins in a spin system). As expected, the COSY spectrum is also dominated by a largely overlap of protons from the cholesterol part with the methyl (CH_3), methylene (CH_2) and allylic (=CHC$\underline{H}_2$) groups of fatty acyl part. MQ spectroscopy appears here as a challenge to search for a specific MQ signature for each molecule based on the separation of molecular fragments through a correlation of their higher quantum coherence order (double quantum (2Q), triple quantum (3Q) and so on) on one dimension along with the regular ^{1}H spectrum (1Q) on the other dimension (Reddy and Caldarelli, 2010, 2011, 2012). Figure 3.5b-e shows the progressive simplification of the correlation peaks for 18:2 EC structure along the spectral MQ series (from 2Q to 5Q), due to the filtering out of spin systems incapable of generating a given MQ coherence. The 2Q spectrum (Figure 3.5b) gives very similar information to that obtained from COSY spectrum (Figure 3.5a), since coupled spins are detected without the diagonal peaks. The C19H_3 (s, δ 1.02) resonance is not detected here since there are no coupled spins (see 18:2 EC molecular structure in Figure 3.5a). In Figure 3.5c, the 3Q spectrum shows lesser peaks when compared with 2Q, namely the C18H_3 (s, δ 0.68), C26H_3 (d, δ 0.86), C27H_3 (d, δ 0.87) and C21H_3 (d, 0.91) resonances. Three of these methyl groups are present in the side chain (C21, C26 and C27) and one (C18) linked to the last ring in cholesterol part. The 4Q spectrum (Figure 3.5d), compared to 3Q, shows a similar pattern of peaks with loss of C9H (m, δ 0.95), C14H (m, δ 0.99) and C4H' (t, δ 2.26) resonances. In addition, some simplification is observed in the region δ 1.21-1.29 corresponding to the methylene protons of the fatty acyl chain. Finally, the 5Q spectrum (Figure 3.5e) shows the simplest spectrum obtained, with loss of resonances in regions δ 1.09-1.20 (C17H, C23H_2, C24H_2, C22H_2, C1H'), δ 1.40-1.50 (C11H, C12H) and δ 1.93-2.02 (C7H, C7H'), when compared with the 4Q spectrum (Figure 3.5d). Hence, the 5Q spectrum of 18:2 EC gives information about the CH_3 (t, δ 0.89), $(CH_2)_n$ (m, δ 1.30), =CHC$\underline{H}_2$ (q, δ 2.04), CH_2CO (m, δ 2.30), =CHC$\underline{H}_2$CH= (t, δ 2.77) and HC=CH (m, δ 5.35) of 18:2 fatty acyl chain, as well as the C1H (m, δ 1.06) and C2H (t, δ 1.58) of cholesterol part. Overall, the MQ spectra showed reduced peak intensities possible due to loss of magnetization during the two τ delays for MQ creation (Keeler, 2002a), especially evident in 5Q spectrum (longer τ delays).

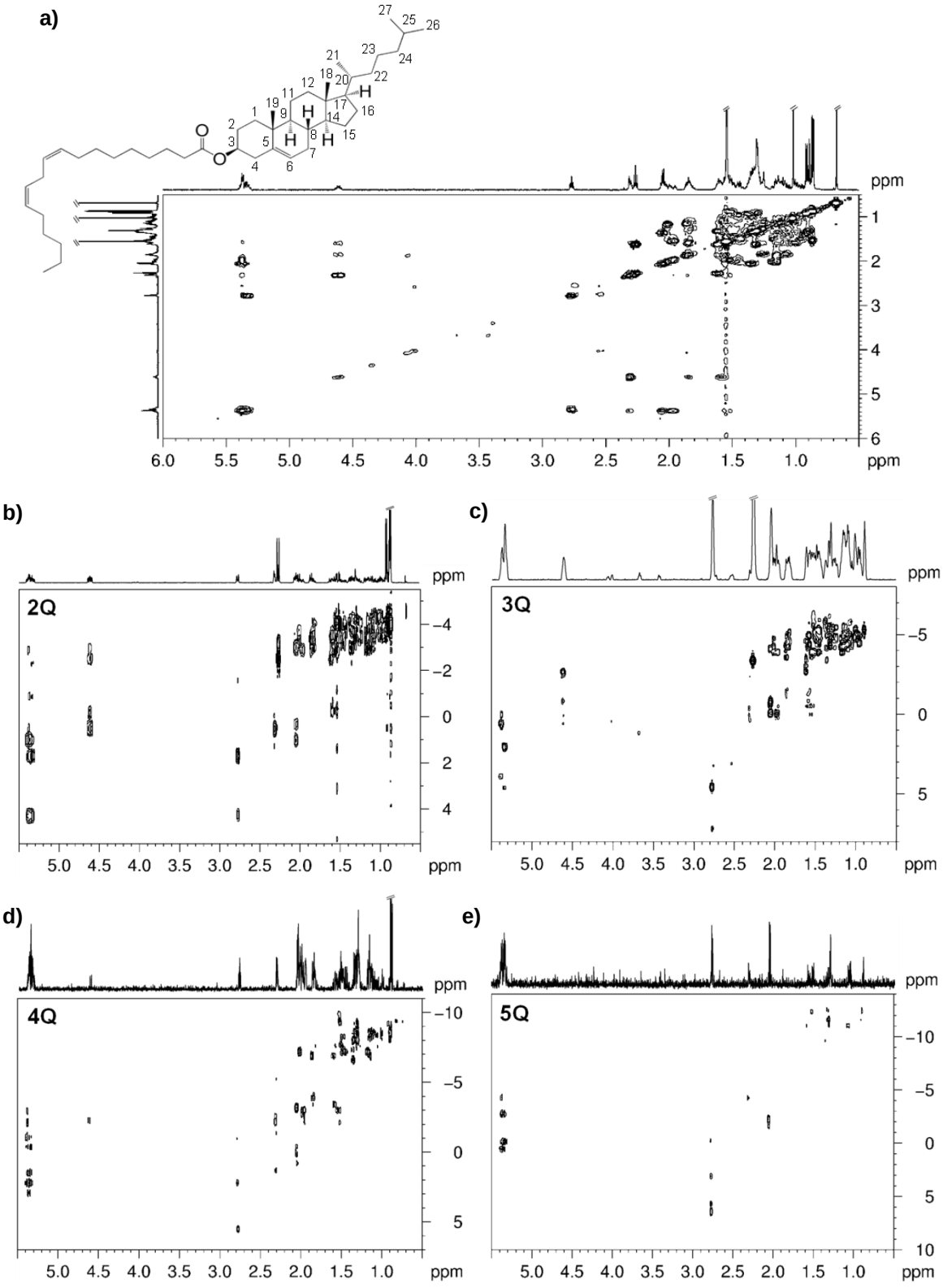


Figure 3.5. 600 MHz a) ^{1}H-^{1}H COSY, b-e) series of MQ-1Q correlation spectra of the 18:2 EC pure standard (1.25 mM), with the coherence order ranging from 2 to 5. The cholesteryl part of 18:2 EC in a) is numbered as originally described by Fieser and Fieser in 1959 (Christie, 2014).

The series MQ-1Q correlation spectra obtained for the lipid mixture showed a fingerprint equal to the sum of the pure compound spectra for the corresponding coherence order. As mentioned above, the 2Q spectrum showed similar information to COSY, where the resonances mentioned above for 18:2 EC (Figure 3.5b) were found with the addition of characteristic resonances from 20:4 FA ($C\underline{H}_2CH_2CO$, $C\underline{H}_2CH=$, CH_2CO, $=CHC\underline{H}_2CH=$) and PL (PE $C\underline{H}_2NH_3$; LPC, PC and SM $N(CH_3)_3$ and $C\underline{H}_2N(CH_3)_3$; LPC, PC and PE glyceryl; $C\underline{H}_2CONH$ SM). These resonances showed large overlap making very hard the identification of LPC proton resonances. The 3Q spectrum of lipid mixture is shown in Figure 3.6a, where it is possible to identify specific resonances from 20:4 FA (A boxes), EC (B boxes), SM (C box), 18:2 FA (D box), and PC+PE (E boxes).

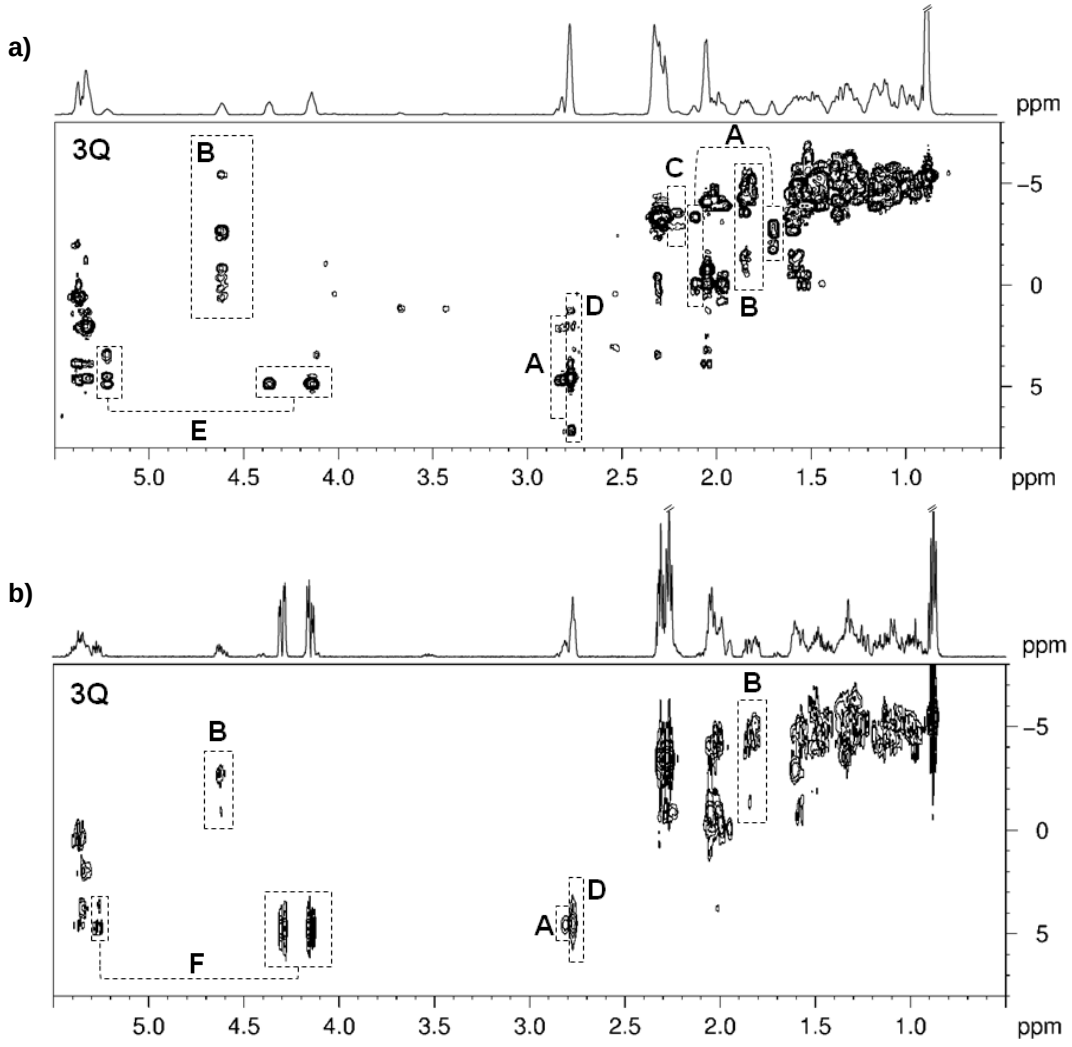


Figure 3.6. a) 600 MHz 3Q-1Q spectrum of lipid mixture (16:0 LPC, 16:0/18:2 PC and PE, d18:1/16:0 SM, 18:2 EC and 20:4 EC). b) 500 MHz 3Q-1Q spectrum of plasma lipid extract of a healthy pregnant woman in 1st T. A: $C\underline{H}_2CH_2CO$ (δ 1.70), $C\underline{H}_2CH=$ (δ 2.11) and $=CHC\underline{H}_2CH=$ (δ 2.82) of 20:4 FA, B: C4H, C15H' and C16H' (δ 1.77-1.90) and C3H (δ 4.61) of cholesterol part of EC, C: $C\underline{H}_2CHNH$ (δ 2.20) of SM, D: $=CHC\underline{H}_2CH=$ (δ 2.77) of 18:2 FA, E: glyceryl protons (δ 4.14, δ 4.84, δ 5.22) of PC+PE, F: glyceryl protons (δ 4.14, δ 4.80, δ 5.26) of TG.

This spectrum gives some simplification compared to 2Q-1Q, mainly in the region δ 2.16-2.24 where now it is possible to identify the C$\underline{H}_2$CONH resonance of SM (C box in Figure 3.6a). Compared to 3Q-1Q, the 4Q-1Q spectrum showed the disappearance of the resonances in the δ 3.30-5.27 and δ 5.42-5.80 regions, corresponding to the glyceryl protons of PC and PE, and unsaturated protons of SM, respectively. Hence, in the 4Q-1Q spectrum, the information about the presence of PL in the mixture is lost. Finally, the 5Q-1Q spectrum showed the simplest profile being very similar to that shown in Figure 3.5e for 18:2 EC, only with the addition of δ 1.70 (C$\underline{H}_2$CH_2CO) and δ 2.30 (C$\underline{H}_2$CO) resonances of 20:4 FA. Hence, this spectrum only gives information about the fatty acyl chains and few resonances of cholesterol part of EC.

Finally, the series of MQ experiments performed before for pure lipid standards and lipid mixture were used in the acquisition of a plasma lipid extract of a healthy pregnant woman in 1st T. Interestingly, the 1D spectrum of plasma lipid extract is very similar to the lipid mixture of pure compounds only with the addition of FC (C19H_3 at δ 1.0, C3H at δ 3.52) and TG (4.14 (glyceryl-C1,3H at δ 4.14, glyceryl-C1,3H' at δ 4.29, glyceryl-C2H at δ 5.26) specific resonances. Hence, the 2Q-1Q spectrum of plasma lipid extract is similar to the 2Q-1Q spectrum of lipid mixture with large overlap and containing more signals, namely the glyceryl of TG (overlapped with that from PL) and FC resonances (overlapped with that from EC). The 3Q-1Q spectrum of plasma lipid extract is shown in Figure 3.6b, where a similar profile is observed when compared with the lipid mixture (Figure 3.6a) with main differences only in the presence of glyceryl protons of TG and the absence of glyceryl protons of PL. The 4Q-1Q spectrum of the plasma lipid extract showed a slightly profile difference when compared with the lipid mixture, namely in the absence of resonances from 20:4 FA, possible due to a lower concentration in plasma lipid extract. Finally, the 5Q-1Q spectrum showed only three resonances with very reduced intensities from cholesterol protons (δ 1.13), $(CH_2)_n$ FA (δ 1.30) and HC=CH FA (δ 5.35).

Overall, MQ spectroscopy showed simpler spectral profiles when compared with 1D and ^{1}H-^{1}H COSY spectra, with specificity based on the *J*-coupling networks. Despite the identification of specific signals from different molecules in a mixture, the assignment of a complete structure was not possible, i.e. the attribution of the side fatty acyl chain linked to the PL or EC characteristic groups. However, the unambiguous identification of poly- and monocyclic aromatic hydrocarbons (Reddy and Caldarelli, 2010) and phenolic molecules in an extract of extra virgin oil (Reddy and Caldarelli, 2011) was previously achieved by MQ spectroscopy showing the potential of these technique for these specific type of molecules.

Chapter 4. Evaluation of plasma stability during handling and storage by NMR metabolomics

The results presented in this chapter were fully published in the article below and are here adapted slightly onto the form of a thesis chapter, for the sake of clarity:

Pinto, J.; Domingues, M. R. M.; Galhano, E.; Pita, C.; Almeida, M. do C.; Carreira, I. M.; Gil, A. M. Human Plasma Stability during Handling and Storage: Impact on NMR Metabolomics. *The Analyst* **2014**, *139*, 1168–1177.

Abstract

This work contributes to fill in some existing gaps in the knowledge of human plasma degradability during handling and storage, a paramount issue in NMR metabolomics. Regarding the comparison between heparin and EDTA anti-coagulant collection tubes, the former showed no interference of the polysaccharide, while conserving full spectral information. In relation to time/temperature conditions, room temperature was seen to have a large impact on lipoproteins and choline compounds from 2.5 hours. In addition, short-term storage at -20^{o}C was found suitable up to 7 days but, for periods up to 1 month, -80^{o}C was recommended. Furthermore, in the case of reusing plasma samples, no more than 3 consecutive freeze-thaw cycles were found advisable. Finally, the impact of long-term -80^{o}C storage (up to 2.5 years) was found almost negligible, as evaluated on a partially matched non-fasting cohort (n=49), after having investigated the possible confounding nature of the particular non-fasting conditions employed.

Brief state of the art

Over the years, much concern has been expressed in relation to the high perishability of biological samples such as blood plasma/serum (Teahan et al., 2006; Barri and Dragsted, 2013) and urine (Dunn et al., 2008; Bernini et al., 2011), and its impact on the detectable metabolome, viewed both by NMR and MS. It has been recognised that the implementation of adequate intra- and inter-laboratory standard operating procedures (SOPs) for blood samples is imperative and several possible bias sources have been investigated 1) blood plasma collection tubes containing different anti-coagulants, namely EDTA, citrate or heparin (Barton et al., 2010; Barri and Dragsted, 2013; Gonzalez-Covarrubias et al., 2013; Yin et al., 2013), 2) different temperature/time conditions during sample handling, analysis and storage (Deprez et al., 2002; Teahan et al., 2006; Barton et al., 2008; Dunn et al., 2008; Trabi et al., 2013), 3) multiple freeze-thaw cycles (Teahan et al., 2006; Fliniaux et al., 2011), when it is necessary to reuse the sample for confirmation or analysis by complementary methods (e.g. MS, metabolite extraction, genetic, proteomic assays). Besides the collection and handling considerations, potential confounders are expected to have important impacts on the blood metabolome e.g. gender (Lawton et al., 2008), age (Lawton et al., 2008; Yu et al., 2012), BMI (Lawton et al., 2008), diet and fasting/non-fasting (Gu et al., 2007; Park et al., 2009; Peré-Trepat et al., 2010), ethnicity (Lawton et al., 2008), co-morbidities or medication. In spite of the high work volume carried out in this context, a few questions remain unanswered thus justifying the work presented in this chapter.

Therefore, this work reports a human plasma NMR metabolomics study which evaluates the changes in the metabolic profile in connection to a) the use of heparin collection tubes, in comparison to EDTA tubes; b) hourly stability at room temperature up to 21 h, with and without sodium azide as preservative; sample stability c) at -20^{o}C and -80^{o}C up to one month and d) during 5 consecutive freeze-thaw cycles. Additionally, the possible confounding effects of non-fasting conditions (2 hours after uncontrolled meal) and of long-term storage at -80^{o}C (up to 2.5 years) will be presented. This work contributes towards a more complete picture of SOPs for NMR metabolomics of human plasma, while providing information about two potentially important confounders (non-fasting conditions at collection and long-term -80^{o}C storage) in a large partially-matched population.

4.1. Effects of heparin and EDTA tubes on plasma spectral profile

Although EDTA and citrate plasma collection tubes have previously been compared in terms of their impact on the ^{1}H NMR spectra of plasma (Barton et al., 2010), to our knowledge, heparin tubes have not been investigated in this context, a general assumption being that heparin (a linear polysaccharide composed of sulphated glucosamine (GlcN), N-acetyl-D-glucosamine (GlcNAc), D-glucuronic acid (GlcA) and L-iduronic acid (IdoA) (Liu et al., 2009; Tovar et al., 2012) does not add significant broad components to the spectra. Indeed, visual comparison of the standard ^{1}H NMR spectra of plasma collected in EDTA and heparin tubes (overlaid in Figure 4.1a) shows almost total superposition of all regions of the spectra, except for the region where the predominant EDTA forms resonate (free and complexed with Ca^{2+}/Mg^{2+}). Multivariate analysis of the spectra obtained for five different controls was performed, after removing the spectral regions where EDTA peaks resonate. PCA of spectra of sample pairs (Figure 4.1b) confirmed the agreement between EDTA- and heparin-collected spectra, the slight deviations noted being smaller than inter-individual variability (which seemed largely determined by lipid content and BMI value, both higher for subject 2). This was confirmed by similar analysis of the CPMG and diffusion-edited spectra and subtraction of heparin- and EDTA-collected sample spectra (not shown). Spectral subtraction unveiled higher levels of pyruvate (26.6 ± 4.7%, singlet at 2.36 ppm, p=0.0023) and two unassigned compounds (111.9 ± 5.7%, singlet at 2.64 ppm, p=0.0079 and 121.1 ± 12.1%, singlet at 3.31 ppm, p=0.0043) in EDTA-collected samples. Pyruvate was indeed detected in a concentrated blank solution left in an

EDTA tube (for 20-30 min), together with acetate (1.91 ppm, s), formate (8.45 ppm, s), lactate (1.33 ppm, d) and several other unidentified small peaks (Figure 4.2a).

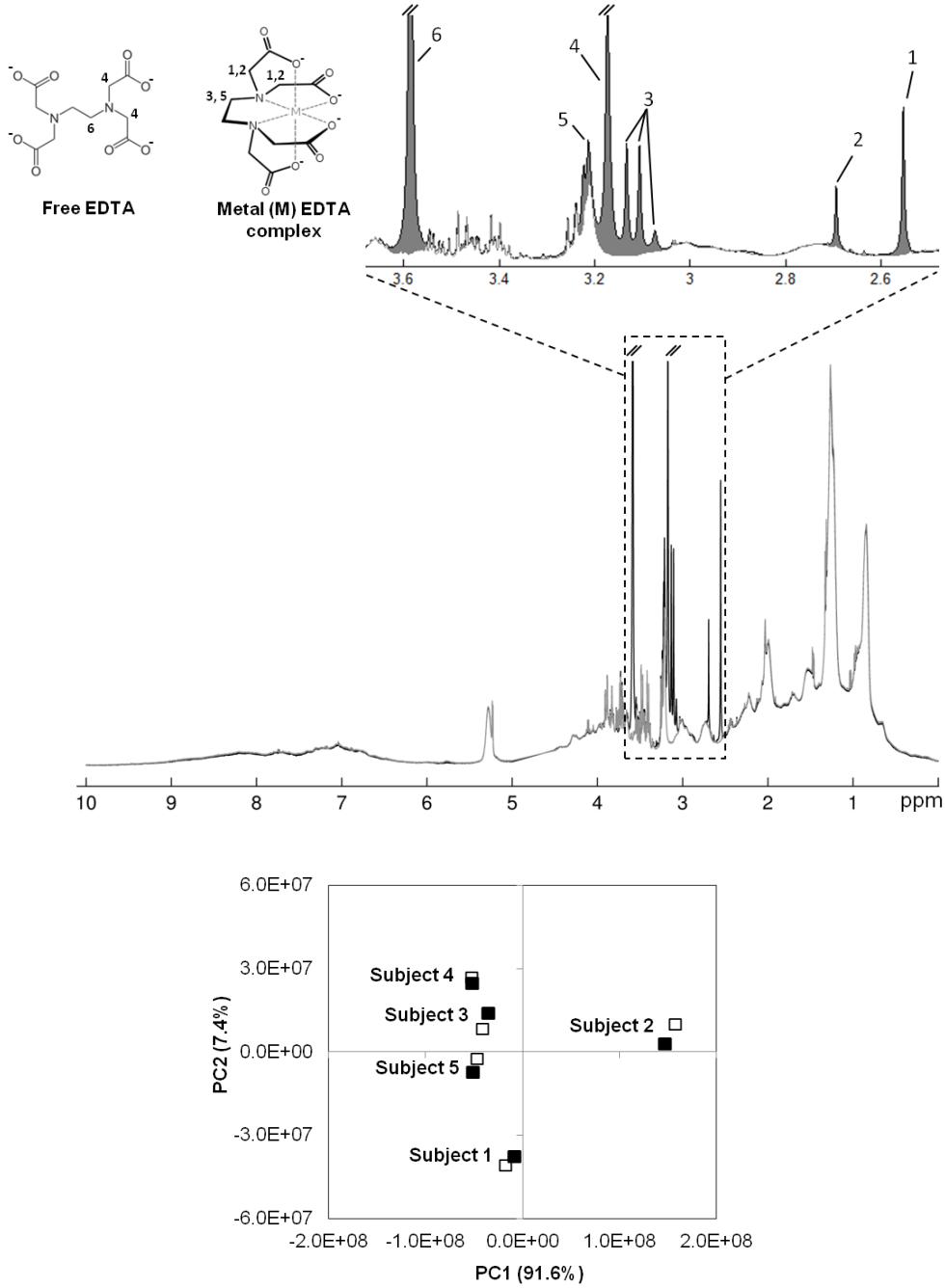


Figure 4.1. a) Superimposed standard ^{1}H NMR spectra of plasma from the same subject, collected with an EDTA tube (black) and a sodium heparin tube (grey). Peak legend: 1- $NC\underline{H}_2CO$ of $CaEDTA^{2-}$, 2- $NC\underline{H}_2CO$ of $MgEDTA^{2-}$, 3- $NC\underline{H}_2CH_2N$ of $CaEDTA^{2-}$, 4- NCH_2CO of free EDTA, 5- NCH_2CH_2N of $MgEDTA^{2-}$ overlapped with $N(CH_3)_3$ of choline-containing metabolites, 6- NCH_2CH_2N of free EDTA (assignment of EDTA resonances based on reference (Barton et al., 2010)); b) PCA scores plot of the standard ^{1}H NMR spectra of plasma collected into EDTA tubes (■) and into sodium heparin tubes (□), after exclusion of the regions accommodating EDTA peaks (2.50-2.60 ppm, 2.66-2.72 ppm, 3.04-3.28 ppm and 3.53-3.65 ppm), filled in grey in Figure 4.1a.

Regarding heparin tubes, the corresponding concentrated blank solution spectrum (Figure 4.2b) showed a clear broad polysaccharide profile, together with a lesser number of interfering resonances (residual formate, lactate and two unidentified signals). For both tube types, however, the expected contaminants concentration should be negligible compared to the plasma spectrum (note that the blank solutions are 10x and 15x more concentrated than a typical plasma sample, Figure 4.2). Due to the overlap with EDTA peaks, we could not confirm the higher choline levels (up to 10%) noted in a previous targeted LC-MS/MS study (Yue et al., 2008) for heparin-collected samples, compared to EDTA ones. Given the above results, both EDTA and heparin tubes seem equally suitable for plasma collection for NMR analysis, as long as the loss of sample resonances overlapping with EDTA peaks is affordable. Overlapped metabolites identified here comprised choline, citrate, dimethylamine, glucose, glycerol, *myo*-inositol, His, Met, Thr, Tyr and Val. Of these, citrate, free choline and dimethylamine do not give rise to other resonances and choline-containing compounds such as PC, LPC and SM are not easily studied in other spectral regions such as the aliphatic or glyceryl regions. A comprehensive list of possible plasma metabolites overlapping with EDTA resonances may be found elsewhere (Barton et al., 2010).

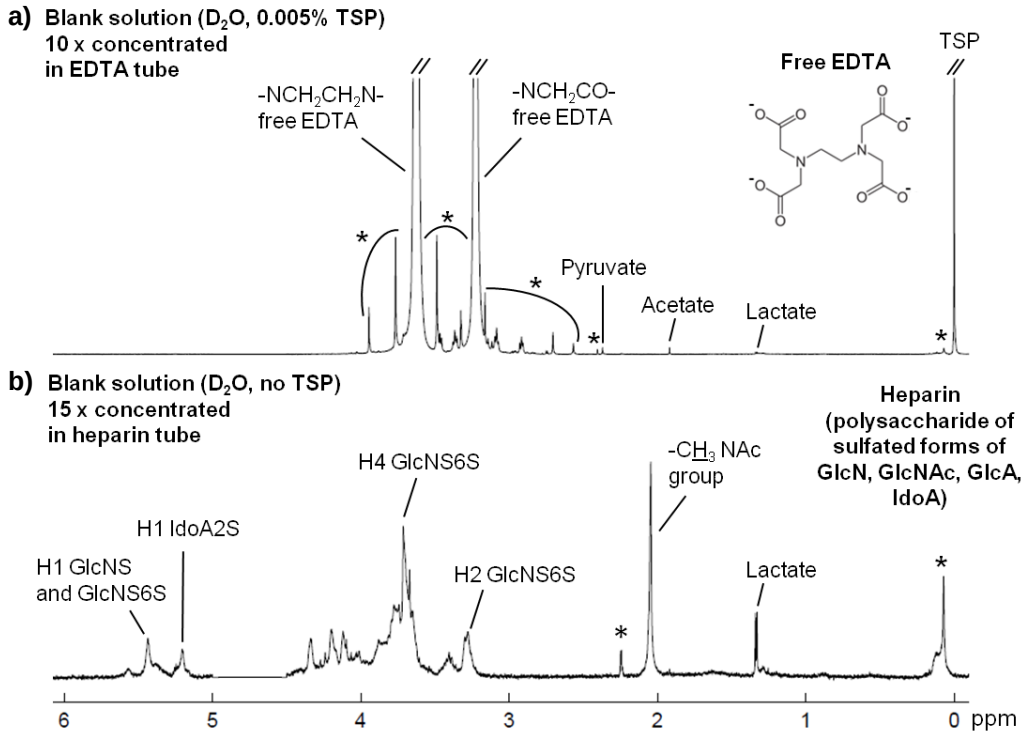


Figure 4.2. Standard ^{1}H NMR spectra of blank solutions added to (a) an EDTA collection tube (D_2O, 0.005% TSP) and (b) a sodium heparin collection tube (D_2O, no TSP). In order to obtain spectra with good signal-to-noise ratio, the blank solutions were concentrated by a factor of 10 and 15 times, respectively (EDTA and heparin assignments based on references (Liu et al., 2009; Barton et al., 2010).

4.2. Short-term plasma stability at room temperature

Considering either the standard or edited (CPMG and diffusion-edited) ^{1}H NMR spectra, the effects of time at room temperature on plasma composition became clear roughly from 2.5 h onwards, as shown below. PCA of the CPMG ^{1}H NMR spectra of three independent plasma samples recorded at room temperature, approximately every hour up to 21 hours, showed that samples follow similar variation trends (Figure 4.3a). It is noted that the plasma of subject 7, incidentally richer in lipids, seems to exhibit a higher magnitude of variation. Computation of the coefficient of variation (Figure 4.3b) indicated that the most significant changes (marked red/orange) affect lipid and choline resonances. Based on this, several resonances were integrated and plotted in a heatmap (Figure 4.3c). This showed that some aliphatic lipid resonances were decreased from *ca.* 2.5 h, including LDL/VLDL $(CH_2)_n$ and CH_3 groups at 1.270 and 0.875 ppm (Liu et al., 2002), respectively. On the other hand, HDL methyls, at 0.850-0.820 ppm (Liu et al., 2002), were significantly increased, reflecting a *ca.* 6% increase in the intensity ratio [CH_3 HDL]/[CH_3 LDL+VLDL]. In addition, several choline resonances were noted to increase, namely the overlapped contributions of PC, LPC, SM (Soininen et al., 2007) at 3.210 ppm and the peak at 3.194 ppm. We suggest that the increase in choline phospholipids may be related to the HDL/(LDL+VLDL) increase, considering the possible discarding of phospholipids from LDL+VLDL (which, taken together, comprise higher phospholipid content) to form HDL. The peak at 3.194 ppm, assigned here mainly to free choline (since it is intensified in CPMG spectra and almost absent in diffusion-edited spectra), showed an increase of up to 20% with time at room temperature. This is in broad agreement with a previous LC-MS/MS report of increase of choline in human plasma during the first hour at room temperature (*ca.* 10%), reaching up to 30% at > 250 min (Yue et al., 2008), due to enzymatic cleavage of choline esters. This process seems to occur concomitantly with the lipoprotein changes, as shown by a 1D STOCSY experiment on the free choline resonance (Figure 4.4a). This confirmed that choline seems to be biochemically correlated to HDL and phospholipids (positive correlation) and to LDL/VLDL (negative correlation). The STOCSY results also helped to identify cleaner and more resolved spin systems for each lipoprotein type: HDL (CH_3: 0.820, 0.840, 0.850; $(CH_2)_n$: 1.220, 1.245; other CH_2: 1.52, 1.98, 2.73; CH=CH: 5.28 ppm) and LDL/VLDL (CH_3: 0.875; $(CH_2)_n$: 1.270; other CH_2: 1.57, 2.01, 2.22; CH=CH: 5.30, 5.33 ppm). In this way, two main HC=CH environments were identified as referring to HDL (5.28 ppm) and LDL/VLDL (5.30 ppm) and, thus, varying in opposite directions with the time at room temperature (Figure 4.4b).

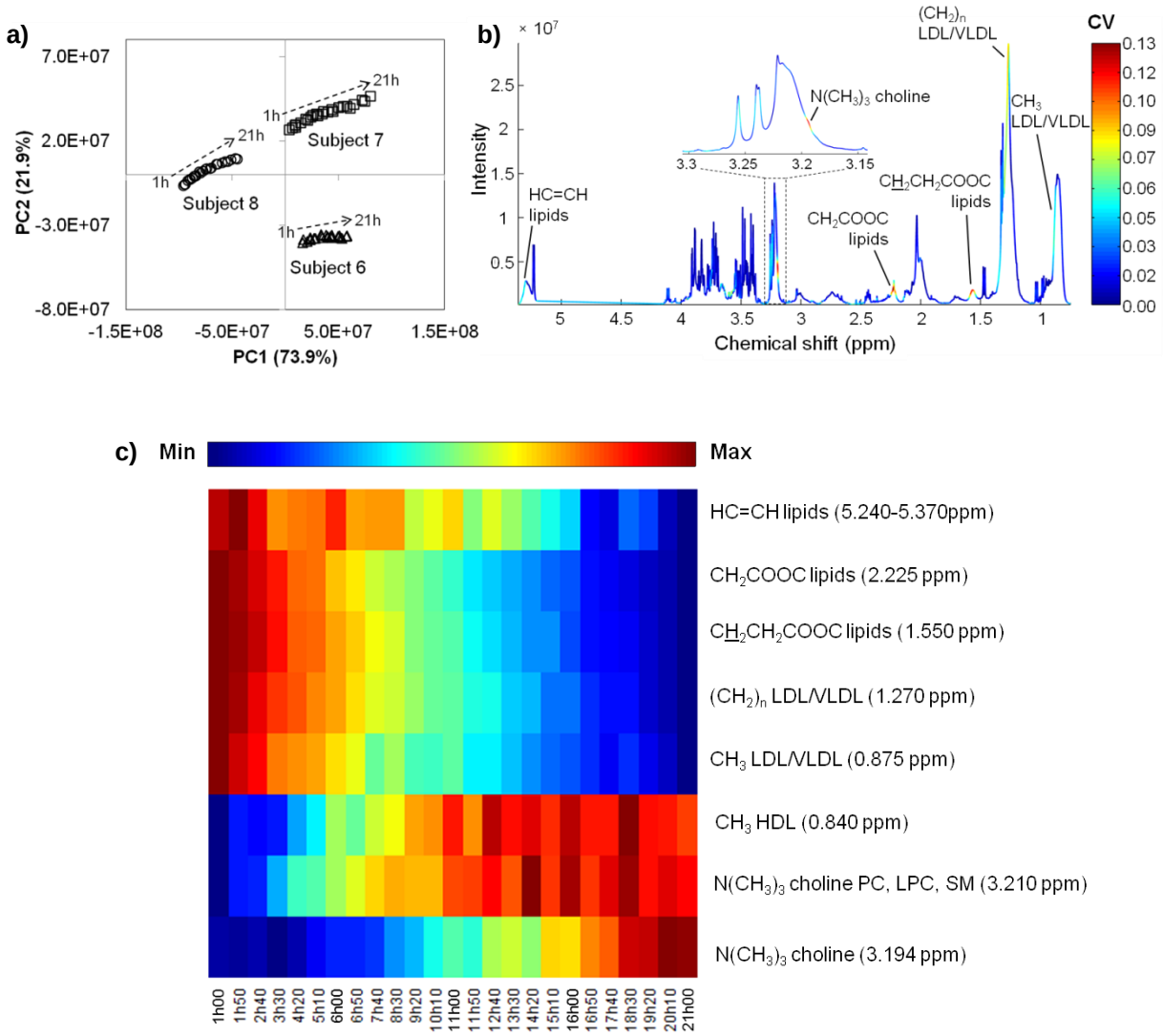


Figure 4.3. a) PCA scores plot obtained for the plasma CPMG ^{1}H NMR spectra of subjects 6, 7 and 8, up to 21h at room temperature; b) average (of three subjects) CPMG ^{1}H NMR spectrum, colored for coefficient of variation (CV) as determined for the overall 21h period; c) heatmap representing the average CPMG integrals of the varying metabolites from minimum (dark blue) to maximum (dark red) values; metabolites are ordered from those exhibiting decreases (above) to those showing increases (below); in the case of largely overlapped regions, the chemical shift indicated corresponds to the maximum intensity observed.

The inverse prominence of these resonances in the CPMG and diffusion-edited spectra (Figure 4.4b) suggested higher mobility of LDL+VLDL unsaturated lipids, compared to those in HDL. Since the average unsaturation degree given by the area ratio [all HC=CH]/[all CH_3] remains unchanged, a conservative change of fatty acid chains from LDL/VLDL to HDL environments seems to be taking place, without lipid oxidation occurring. As to the origin of the above changes, the room temperature study carried out with sodium azide produced identical changes (not shown) as in its absence, indicating the occurrence of enzymatic lipolytic action without microbial growth (also

confirmed by the absence of other indicators of microbial growth e.g. lactate). Previous studies have indeed reported changes in lipids after 3 hours (Teahan et al., 2006) and 6 hours (Bernini et al., 2011) at room temperature, however in not as much detail as in this study. Other changes reported in relation to amino acids, glycerol and citrate (Deprez et al., 2002; Teahan et al., 2006; Bernini et al., 2011) were not noted here.

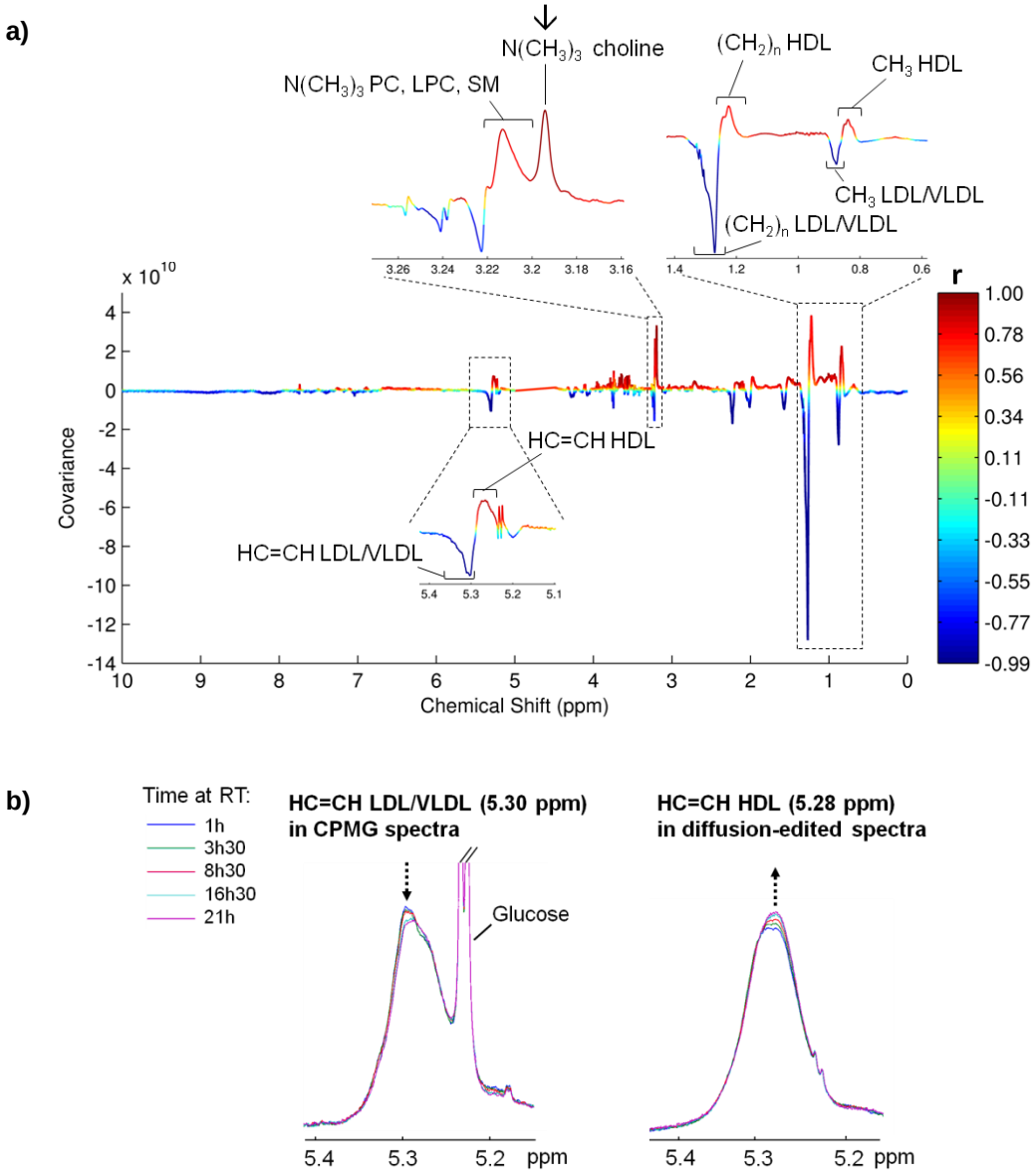


Figure 4.4. a) 1D STOCSY obtained using the -$N(CH_3)_3$ choline peak at 3.194 ppm as root peak (see arrow in inset), with color scale expressing the correlation (r) value; b) expansions of the HC=CH region of the average CPMG (left) and diffusion-edited (right) ^{1}H NMR spectra recorded as a function of time, at room temperature (only selected times are shown). RT: room temperature.

4.4. Short-term plasma stability at -20ºC and -80ºC and effect of freeze-thaw cycles

Regarding the compositional stability of plasma stored up to 1 month at -20°C and at -80°C, both standard and diffusion-edited spectra showed remarkable spectral agreement over the whole period (not shown), only the CPMG spectra reflecting the occurrence of small variations, as shown by PCA (Figure 4.5a). Visual spectral inspection and integration confirmed small changes observable only at 31 days (Figure 4.5b): 1) proline (*ca.* +31%, p=0.033, at -20°C), 2) glucose (*ca.* +9%, p=0.019, at -20°C) and 3) unassigned broad peak at 6.8-7.1 ppm (*ca.* -50%, p=0.00036 and 0.0022, at -20°C and -80°C respectively), the latter variation probably due to protein precipitation (also supported by visual observation). In previous studies, no significant changes had been seen in the NMR spectra of rat plasma stored for 1 month at -20°C or -80°C (Deprez et al., 2002).

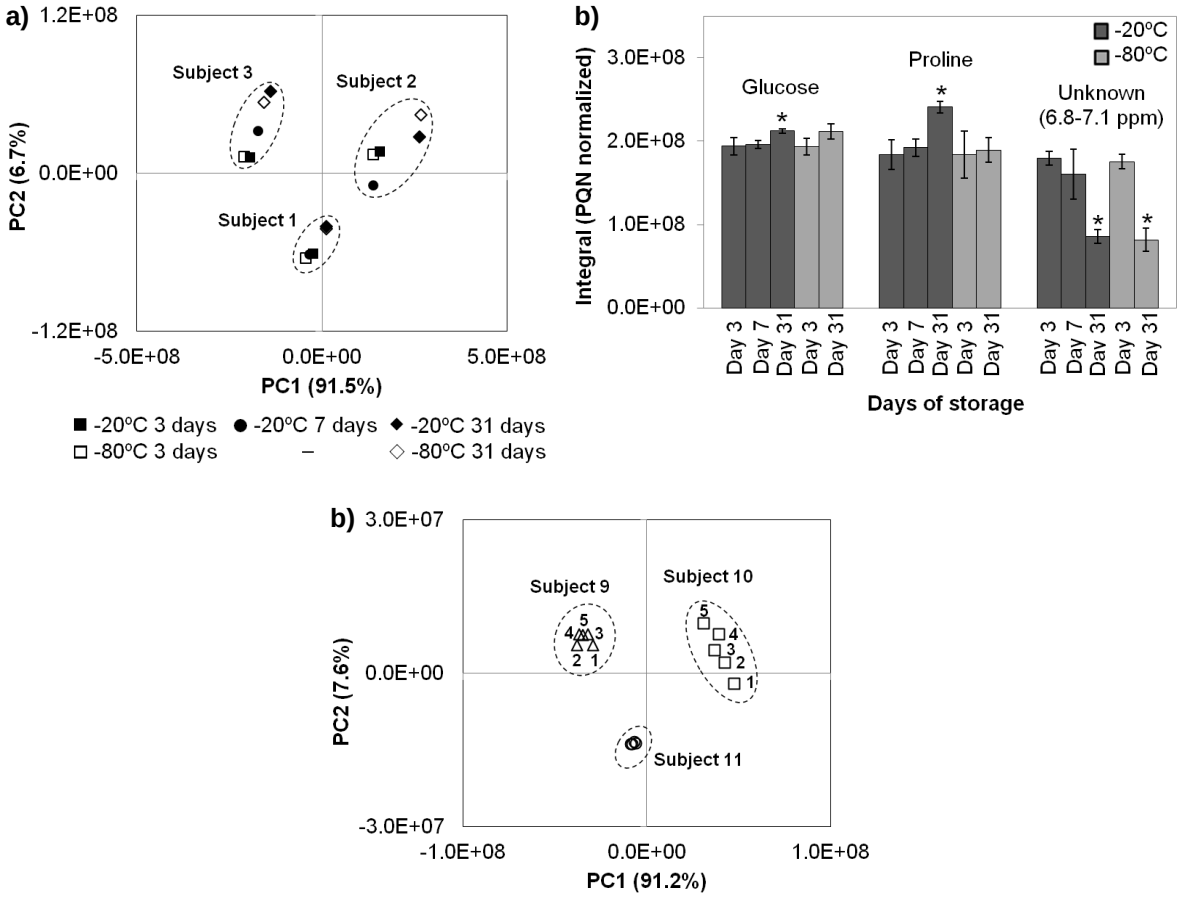


Figure 4.5. a) PCA scores plot obtained for the CPMG ^{1}H NMR spectra of three plasma samples (subject 1, 2 and 3) stored at -20ºC (filled symbols) or at -80ºC (open symbols) for up to 31 days; b) histogram of metabolite integrals illustrating the variations occurring during plasma storage at -20ºC and -80ºC up to 1 month (*, p-values < 0.05); c) PCA scores plot obtained for the standard ^{1}H NMR spectra of three plasma samples (subjects 9, 10 and 11) subjected to consecutive freeze-thaw cycles (numbers indicate number of cycles).

Regarding the effects of up to 5 freeze-thaw cycles on plasma composition, the PCA of standard spectra shown in Figure 4.5c expresses an interesting result, since the degree of sample dispersion resulting from freeze-thaw cycles 1 to 5 is clearly sample dependent. The original plasma composition of subject 10 comprised higher contents of lipids (CH_3, $(CH_2)_n$), choline-containing compounds, Val , Ala and lactate, compared to the remaining two samples, such composition (particularly in regard to lipids) possibly leading to higher degradability of the sample. For this subject, changes upon 5 consecutive freeze-thaw cycles were found to arise from: 1) lipids decrease (*ca.* -5%), 2) choline phospholipids increase (*ca.* 7%), 3) Ala, glucose and pyruvate increases (*ca.* 2-7%) and acetone decrease (*ca.* -16%). These results are in broad agreement with previous studies of 1 freeze-thaw cycle (Teahan et al., 2006), which reported unspecified alterations in lipids, Ala, glucose and lactate. A more recent study (Fliniaux et al., 2011) on serum noted statistically relevant (p value < 0.05) changes upon 5 and 10 freeze-thaw cycles (decreases in choline resonance at 3.20 ppm, glycerol, methanol, ethanol, proline, unassigned peak at 1.91 ppm), none of which having been noted in the conditions of this study. Furthermore, the present study has shown that most variations take place at the fourth freeze-thaw cycles and thereafter, so that no more than 3 freeze-thaw cycles are recommendable.

4.5. Possible confounding factors: subject non-fasting and long-term plasma storage at -80ºC

A cohort of healthy pregnant women in their 2nd trimester of pregnancy, partially matched for age, BMI and gestational age, was considered in order to study the effects of -80^{o}C storage for up to 2.5 years. As such samples were collected in non-fasting conditions, a possible confounder superimposed on the eventual effects of long-term storage, the effects of non-fasting were firstly studied on a group of non-pregnant women. The PCA plot shown in Figure 4.6a indicated that, as expected, each subject responded differently to meal intake (see arrows shown as examples); however, no group separation took place, the overall sample groups remaining largely overlapped. Hence, the particular non-fasting conditions characterising sample collection for the pregnant women in this study (2 hours after ingestion of uncontrolled breakfast) are not expected to have a strong effect on the overall characteristics of the sample group. Regarding long-term storage at -80^{o}C, multivariate analysis of the spectra for 6-12 months and 14-19 months groups (not shown) detected no significant changes, minor changes being noted when comparing the 14-19 months to 20-30 months groups. This was expressed by a weak non-predictive PLS-DA model (Q^2 0.28) (Figure 4.6b), which

upon inspection of the spectra translated into a small (*ca.* 2%) increase in cholesterol ($p = 0.023$) and slight variations in N-acetyl glycoproteins and creatine (found to be largely determined by two outliers with larger BMI and gestational age, respectively), in the 20-30 months storage period.

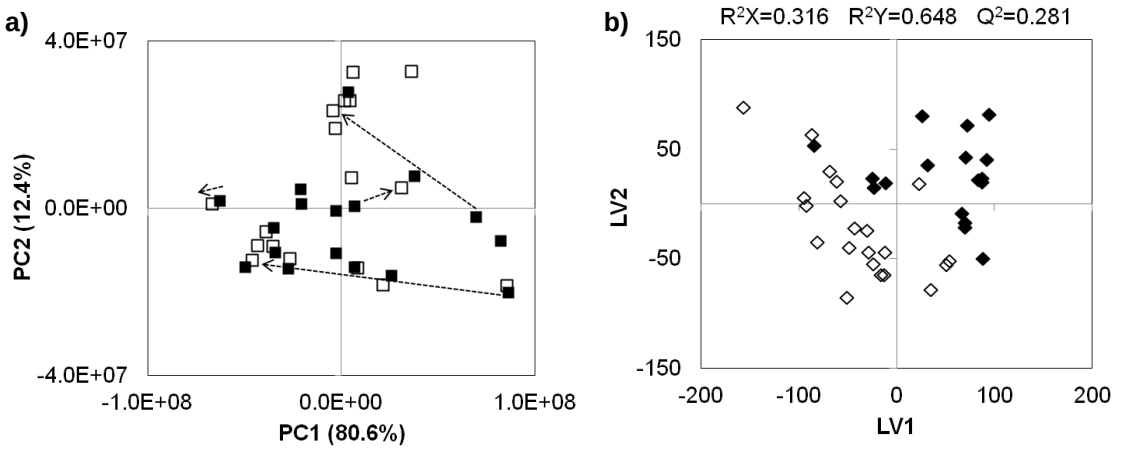


Figure 4.6. a) PCA scores plot obtained for the standard ^{1}H NMR spectra of plasma samples obtained for the same set of subjects (n=16), under fasting (■) and non-fasting (□) conditions; arrows represent changes from fasting to non-fasting for selected subjects. b) PLS-DA scores plot obtained for standard ^{1}H NMR spectra of plasma samples stored at -80^{o}C for 14-19 months (◆, n=17) and 20-30 months (◇, n=21).

The above results expressed the low degree of change induced on plasma composition, as viewed by NMR, by the non-fasting conditions used in this work and by storage up to 30 months at -80^{o}C, after which samples may safely be studied (although interpreting with care eventual changes in cholesterol, after 20 months). Previous NMR results have shown changes in amino acids resonances and the disappearance of citrate resonances in rat plasma stored at -80ºC for 9 months (Deprez et al., 2002), whereas -20^{o}C storage of bovine plasma for between 2-15 years revealed changes in lipids and several low M_W metabolites (glycerol, β-hydroxybutyrate, amino acids) (Trabi et al., 2013). To our knowledge, no long-term storage study of human plasma had been carried out before by NMR, a very recent report existing on the subject, using UPLC-MS (Hebels et al., 2013), and reporting no significant changes.

4.6. Conclusions

In this work, we have shown that EDTA and heparin tubes seem equally suitable for plasma collection for NMR analysis, as long as 1) the level of interfering peaks (in higher number in EDTA tubes) is kept low compared to sample peaks of interest and 2) the loss of sample resonances overlapping with EDTA peaks (and removed from the

dataset) is affordable. The latter is particularly important for compounds not giving rise to other resonances (citrate, free choline and dimethylamine) and choline-containing compounds PC, LPC and SM, which are not easily studied in other spectral regions such as the aliphatic or glyceryl regions.

Room temperature stability studies revealed significant changes in the lipidic components of plasma, after *ca.* 2.5 hours. Such changes indicated choline formation (increased up to 20%) through enzymatic cleavage of phospholipids, along with LDL/VLDL conversion into HDL (increase of *ca.* 6 % in HDL/(LDL+VLDL) ratio), possibly also in connection with the increase noted in PC, LPC and SM levels. In addition to these compounds, no other significant changes in plasma composition were observed at room temperature (up to 21 hours). In relation to short-term (up to 1 month) storage stability, results revealed that plasma samples may be kept at least up to 7 days at -20^{o}C without significant changes occurring. Upon 1 month of storage, increases in proline and glucose and a decrease of a broad resonance at 6.8-7.1 ppm (possibly reflecting protein precipitation) were observed for samples kept at -20^{o}C. For samples stored at -80^{o}C for the same length of time, only the latter variation was observed. In relation to freeze-thaw cycles (-80^{o}C), the effects of up to 5 cycles were found to be strongly sample dependent, a larger impact noted for a sample richer in lipids. Most variations were seen to take place upon cycle 4 and thereafter, so that no more than 3 freeze-thaw cycles are recommendable. Changes comprised small decreases in lipids and acetone and increases in choline compounds, Ala, glucose and pyruvate. The longer-term stability of human plasma, at -80^{o}C (up to 2.5 years), was studied on a large cohort (n=49) of pregnant women, for whom plasma collection was performed under non-fasting conditions. In order to investigate the possible confounding effect of non-fasting, this factor was singled out first and studied on a smaller group of subjects (n=16), proving not to affect the overall characteristics of the group of samples, as viewed by multivariate analysis. Subsequently, the effects of long-term -80^{o}C storage were found almost negligible up to 30 months, a small increase in cholesterol having been noted after 20 months storage.

Chapter 5. Following healthy pregnancy metabolism by NMR metabolomics of maternal plasma

The results presented in this chapter were fully published in the article below and are here adapted slightly onto the form of a thesis chapter, for the sake of clarity:

Pinto, J.; Barros, A. S.; Domingues, M. R. M.; Goodfellow, B. J.; Galhano, E.; Pita, C.; Almeida, M. C.; Carreira, I. M.; Gil, A. M. Following Healthy Pregnancy by NMR Metabolomics of Plasma and Correlation to Urine. *Journal of Proteome Research* **2014**, *14*, 1263-1274.

Abstract

This work presents the first NMR metabolomics study of maternal plasma during pregnancy, including correlation between plasma and urine metabolites. The expected decrease in circulating amino acids early in pregnancy was confirmed with six amino acids being identified as required by the fetus in larger extents. Newly observed changes in citrate, lactate and dimethyl sulfone suggested early adjustments in energy and gut microflora metabolisms. Alterations in creatine levels were also noted, in addition to creatinine variations reflecting alterations in glomerular filtration rate. Regarding plasma macromolecules, HDL and LDL+VLDL levels were confirmed to increase throughout pregnancy, although at different rates and accompanied by increases in fatty acid chain length and degree of unsaturation. Correlation studies suggested a) an inverse relationship between lipoproteins (HDL and LDL+VLDL) and albumin, with a possible direct correlation to excreted (unassigned) pregnancy markers resonating at δ 0.55 and δ 0.63; b) a direct link between LDL+VLDL and N-acetyl-glycoproteins, together with excreted marker at δ 0.55 and c) correlation of plasma albumin with particular circulating and excreted metabolites. These results have unveiled specific lipoprotein/protein metabolic aspects of pregnancy with impact on the excreted metabolome and, therefore, provide an interesting lead for the further understanding of pregnancy metabolism.

Brief state of the art

In order to define robust prenatal disease biomarkers, the complex metabolic evolutions related to gestational age in healthy pregnancies must be taken into account. Healthy pregnancies have been studied by metabolomics of amniotic fluid (Cohn et al., 2009) and maternal urine (Diaz et al., 2013a; Trivedi and Iles, 2014a) by ^{1}H NMR spectroscopy and by liquid chromatography (LC)-MS, and of maternal blood plasma by LC-MS (Luan et al., 2014). NMR metabolomics has revealed gestational age-related alterations in amniotic fluid levels of alanine, valine, creatinine and glucose (Cohn et al., 2009) and significant changes in 21 maternal urine metabolites (8 of which newly detected in relation to pregnancy: choline, creatinine, 4-deoxyerythronic acid, 4-deoxythreonic acid, furoylglycine, guanidoacetate, β-hydroxybutyrate and lactate) (Diaz et al., 2013a). In addition, a recent untargeted LC-MS and shotgun lipidomics of maternal plasma extracts revealed increases in dihydrobiopterin, free carnitine, acetyl-alanine, leucyl-phenylalanine, LPC (18:1) and LPC (20:4), along with changes in total

concentrations of LPC, PC, lysophosphatidylethanolamines (LPE) and lysophosphatidylserines (LPS) (Luan et al., 2014).

The present paper describes for the first time, to our knowledge, an untargeted ^{1}H NMR study of maternal blood plasma collected for healthy non-pregnant and pregnant women in their 1st, 2nd and 3rd trimesters of pregnancy, as well as at delivery, with the aim of evaluating the metabolic adaptations reflected in plasma throughout healthy pregnancies. This complements the recent MS-based studies (Luan et al., 2014) by giving an overview of the changes in the set of more abundant (mM and μM) plasma metabolites. Furthermore, dynamic changes measured in plasma are correlated here to those observed in maternal urine, a novel approach enabling a more complete two-biofluid metabolic picture of pregnancy to be obtained. The results provide confirmation of known metabolic alterations occurring during pregnancy and reveal additional metabolite changes and relationships, particularly in relation to lipoproteins composition and their links to other circulating and excreted compounds. The detailed metabolic signature of healthy pregnancies obtained here constitutes an important step towards the identification of disease-related deviations allowing improved disease prediction, diagnosis and monitoring.

5.1. Analysis of whole plasma

The average CPMG spectra of maternal plasma across pregnancy (Figure 5.1) show some changes, namely in lipids, choline metabolites, lactate, valine, alanine, glutamine, citrate, creatine, creatinine, histidine and tyrosine. Such changes were more clearly illustrated in the CPMG spectra, compared to the standard and diffusion-edited spectra of blood plasma, where large overlap of broad resonances (arising from high M_w compounds) hinders spectral inspection. PCA of the CPMG ^{1}H NMR spectra of maternal plasma (Figure 5.2A, left) revealed three partially overlapped groups of samples: NP (□), 1st and 2nd T samples (◆ and ○), and 3rd T together with delivery samples (▲ and ∗). The 3rd T outlier corresponds to higher gestational age and BMI, its spectrum reflecting higher lipid and N-acetyl glycoprotein (2.0 ppm, s) contents. PLS-DA (Figure 5.1A, right) unveiled a clearer metabolic trajectory for healthy pregnancies, with overlap of plasma collected in the 3rd T and at delivery. Pairwise PCA (not shown) and PLS-DA (Figure 5.2B) confirmed separation for NP/1st T (R^2X 0.20, R^2Y 0.93, Q^2 0.76, 98% sens., 89% spec.), 1st/2nd T (R^2X 0.18, R^2Y 0.83, Q^2 0.55, 83% sens., 87% spec.) and 2nd/3rd T (R^2X 0.18, R^2Y 0.91, Q^2 0.45, 71% sens., 100% spec.), again with the exception of 3rdT/delivery samples, which overlap in all models. Model robustness was assessed by MCCV and confirmed for all three models (Figure 5.3).

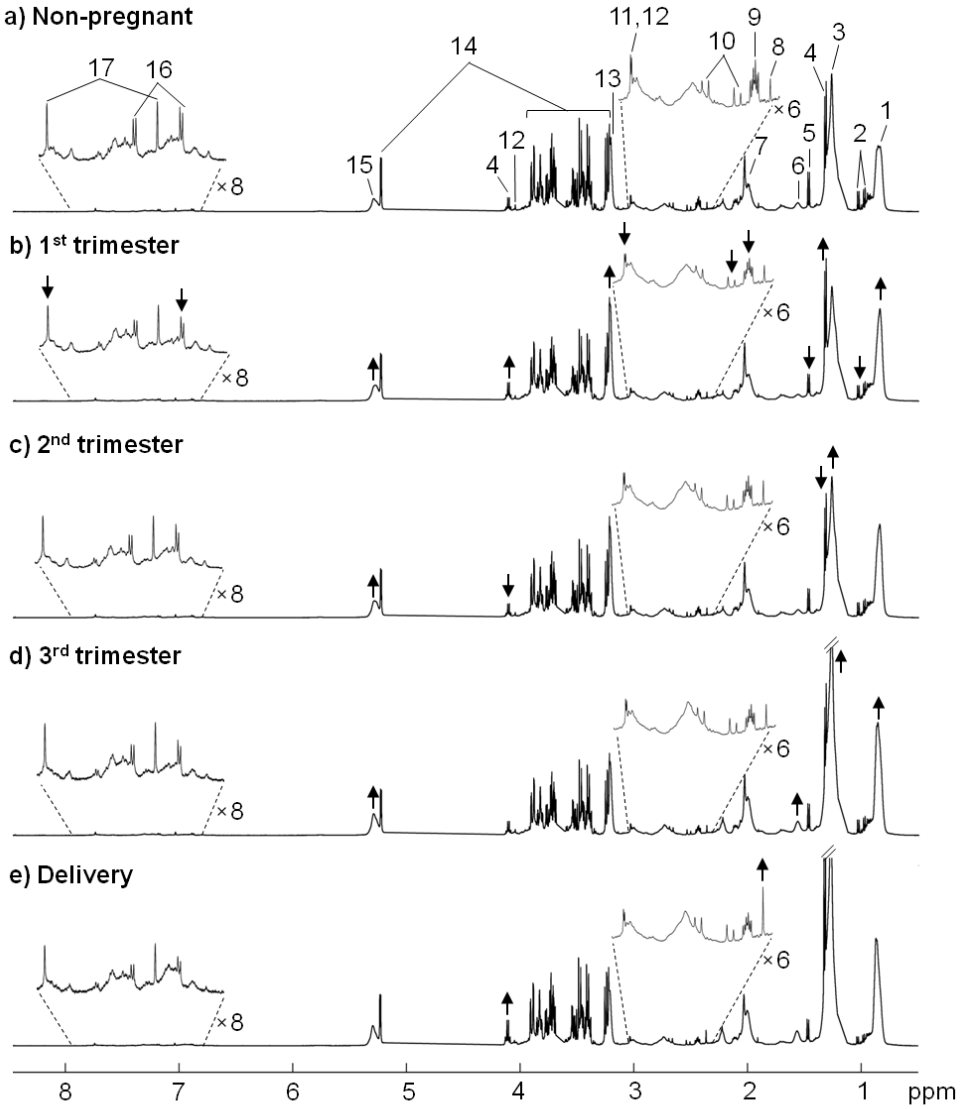


Figure 5.1. Average 500 MHz CPMG ^{1}H NMR spectra of blood plasma of a) non-pregnant (NP) and pregnant women at b) 1st (11-13 g.w.), c) 2nd (16-22 g.w.), d) 3rd (31-40 g.w.) and e) delivery (32-41 g.w.). Assignment (with 3-letter codes for amino acids): 1- CH_3 lipids, 2- Val, 3- $(CH_2)_n$ lipids, 4- lactate, 5- Ala, 6- $C\underline{H}_2CH_2CO$ lipids, 7- N-acetyl glycoproteins, 8- pyruvate, 9- Gln, 10- citrate, 11- creatine, 12- creatinine, 13- $N(CH_3)_3$ choline metabolites, 14- glucose, 15- HC=CH lipids, 16- Tyr, 17- His. The arrows guide the eye for some visible spectral alterations (increase and decrease).

Based on the analysis of loading plots (Figure 5.2B, right) and peak integration, 11 low M_W metabolites (6 amino acids, citrate, creatine, creatinine, dimethyl sulfone and lactate) were found to vary significantly at some stage in pregnancy, together with lipoproteins (HDL, LDL+VLDL) and other proteins (N-acetyl glycoproteins δ 2.0, albumin-lysyl δ 2.9-3.0 and the set of broad NH resonances reflecting total protein δ 5.5-9.0) (Table 5.1).

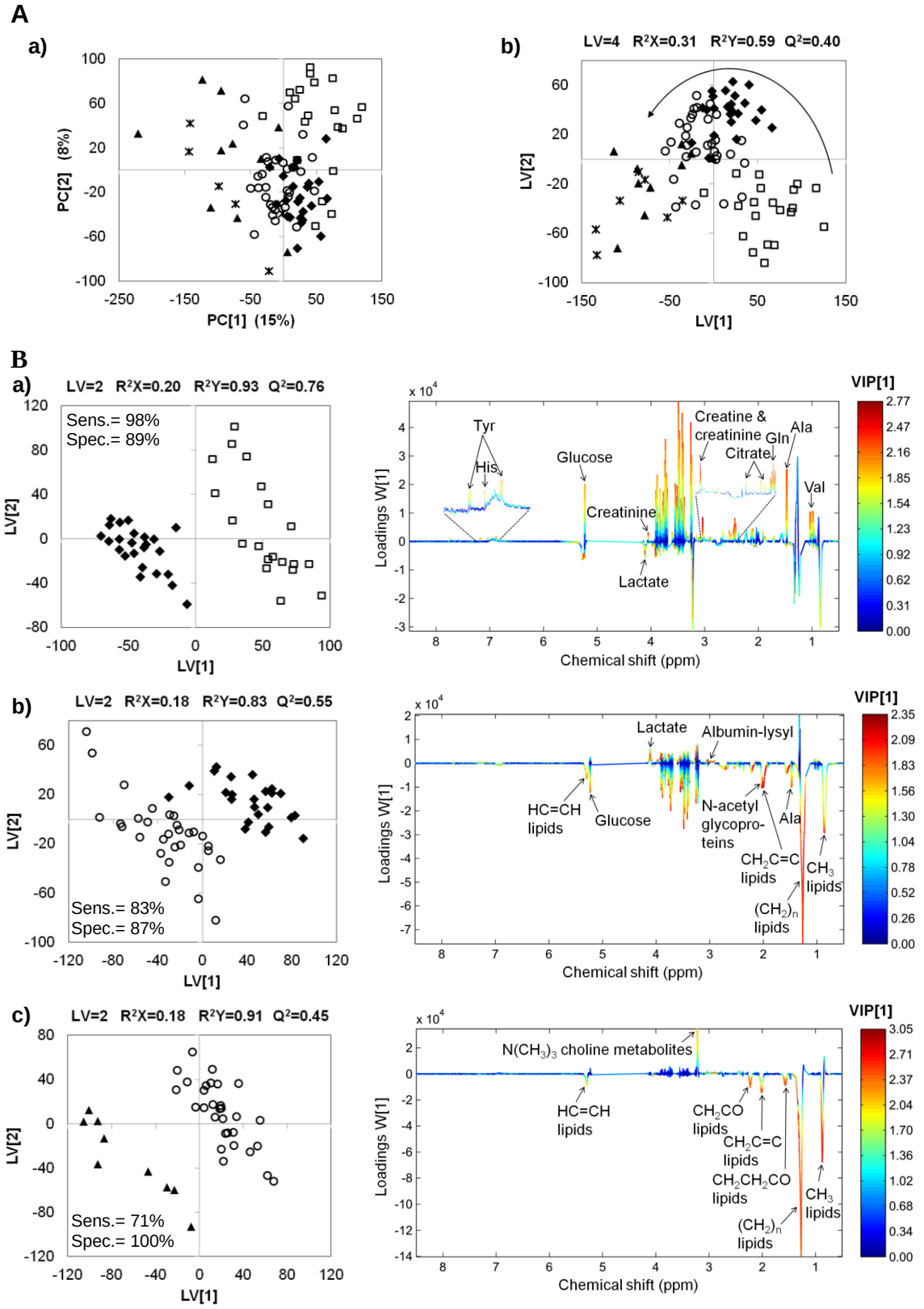


Figure 5.2. A) PCA a) and PLS-DA b) scores scatter plot for CPMG ^{1}H NMR plasma spectra of NP (□), and pregnant women at 1st T (◆), 2nd T (○), 3rd T (▲) and delivery (∗). The arrow in a) indicates a 3rd T outlier with higher BMI than the remaining samples. The curved arrow in b) indicates the trajectory followed across pregnancy. B) PLS-DA score (left) and loading plots (right) for CPMG ^{1}H NMR plasma spectra of a) NP (□) vs 1st T (◆), b) 1st T (◆) vs 2nd T (○) and c) 2nd T (○) vs 3rd T (▲) groups. Loadings are colored according to VIP. Three-letter codes are used for amino acids.

Changes in citrate, creatine, dimethyl sulfone and lactate were registered here for the first time, to our knowledge, related to pregnancy (* in Table 5.1). The information shown in Table 5.1 and time course changes (selected examples in Figure 5.4) indicate that relevant changes in low M_W metabolites occur mainly from the NP state to the 1st T, with decreased levels of valine, alanine, glutamine, tyrosine, histidine and citrulline, citrate, creatine, creatinine and dimethyl sulfone, and a suggestion of increased lactate levels.

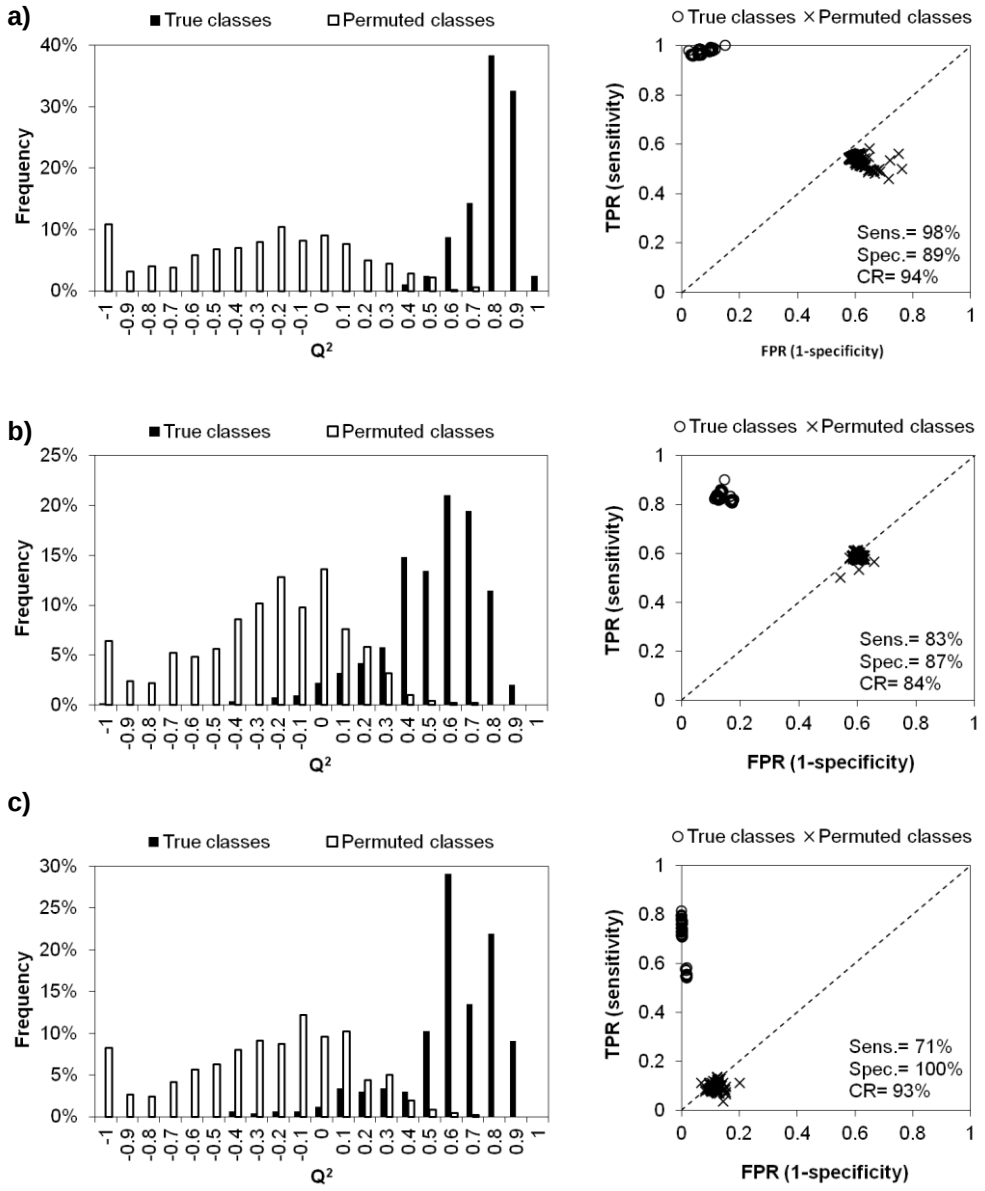


Figure 5.3. True and permuted model Q^2 distributions (left) and ROC plots (right) given by MCCV of the PLS-DA models obtained for CPMG ^{1}H NMR plasma spectra of a) NP vs 1st T, b) 1st T vs 2nd T and c) 2nd T vs 3rd T groups. TPR: true positive rate, FPR: false positive rate.

Table 5.1. List of blood plasma compounds changing in content from the NP state to 1st T, from 1st to 2nd T, and from 2nd to 3rd T. s: singlet, d: doublet, q: quartet, m: multiplet, br: broad signal. [a] %Variation determined as $(\bar{X}_a - \bar{X}_b)/\bar{X}_b \times 100$, where $\bar{X}$ denotes the average value for each group *a* and *b*. [b] E.S.: effect size determined as described in (Berben et al., 2012); values in brackets correspond to high uncertainties. [c] Significance level 95% (p-value < 0.05). [d,e,f] Metabolite integrals measured in the CPMG, diffusion-edited and standard spectra, respectively. [*] Metabolites observed to vary in maternal blood plasma for the first time throughout pregnancy.

Compound	Chemical shift/ppm	NP *vs.* 1st T			1st *vs.* 2nd T			2nd *vs.* 3rd T		
		%Variat.[a]	E.S.[b]	p-value[c]	%Variat.[a]	E.S.[b]	p-value[c]	%Variat.[a]	E.S.[b]	p-value[c]
Low molecular weight										
Alanine[d]	1.45 (d)	-17.7±5.1	-1.3±0.6	6.7×10^{-4}	7.4±3.0	0.6±0.5	2.2×10^{-2}			
Glutamine[d]	2.45 (m)	-20.1±4.6	-1.6±0.7	4.2×10^{-5}				-10.7±3.4	-1.1±0.8	4.2×10^{-3}
Histidine[d]	7.04 (s)	-11.4±4.8	-0.8±0.6	1.8×10^{-2}						
Tyrosine[d]	6.89 (d)	-30.9±11.8	-1.0±0.6	1.1×10^{-4}	22.4±5.7	0.9±0.6	1.2×10^{-3}			
Valine[d]	1.03 (d)	-8.8±3.8	-0.7±0.6	2.1×10^{-2}						
Citrate[d,*]	2.51 (m)	-12.7±4.8	-0.9±0.6	2.9×10^{-3}	5.6±3.3	(0.4±0.5)	4.0×10^{-2}			
Citrulline[d]	1.88 (m)	-14.4±2.8	-1.8±0.7	6.4×10^{-6}						
Creatine[d,*]	3.03 (s)	-18.7±4.5	-1.5±0.7	5.3×10^{-5}						
Creatinine[d]	4.05 (s)	-14.1±3.1	-1.6±0.7	3.0×10^{-5}				14.6±5.2	1.1±0.8	2.0×10^{-2}
Dimethyl sulfone[d,*]	3.14 (s)	-19.4±4.8	-1.5±0.7	1.1×10^{-4}						
Lactate[d,*]	4.1 (q)	35.5 ± 8.8	0.9 ± 0.6	4.7×10^{-4}	-21.4±9.0	-0.8±0.6	1.4×10^{-3}			
High molecular weight										
Lipid resonances:										
C18H cholest.[e]	0.6-0.7 (br)							5.2±1.5	1.5±0.8	5.5×10^{-3}
CH_3[e]	0.8-0.9 (br)	5.1±2.4	(0.6±0.6)	4.3×10^{-2}	7.9±1.8	1.1±0.6	1.5×10^{-4}	21.7±4.3	2.6±0.9	7.2×10^{-4}
$(CH_2)_n$[e]	1.1-1.4 (br)	7.4±3.3	0.7 ±0.6	3.8×10^{-2}	15.0±2.5	1.5±0.6	6.7×10^{-7}	34.7±6.9	2.5±0.9	1.0×10^{-3}
$C\underline{H}_2CH_2CO$[e]	1.4-1.6 (br)				4.9±0.9	1.4±0.6	1.8×10^{-6}	12.1±2.6	2.4±0.9	1.0×10^{-3}
$C\underline{H}_2CH=$[e]	1.9-2.0 (br)				6.6±1.2	1.4±0.6	3.2×10^{-6}	15.2±3.0	2.6±0.9	5.9×10^{-4}
CH_2CO[e]	2.2-2.3 (br)				7.2±1.3	1.4±0.6	1.3×10^{-6}	18.6±4.4	2.2±0.9	2.3×10^{-3}
$=CHC\underline{H}_2CH=$[e]	2.6-2.8 (br)	7.1±2.1	1.0±0.6	3.7×10^{-4}	5.5±1.4	1.1±0.6	2.4×10^{-4}	6.6±2.2	1.2±0.8	1.2×10^{-2}
$N(CH_3)_3$ cho.[e]	3.2-3.3 (br)	12.8±3.0	1.2±0.6	3.5×10^{-4}	5.6±2.3	0.6±0.5	2.2×10^{-2}	8.6±2.9	1.1±0.8	1.1×10^{-2}
$C\underline{H}_2N(CH_3)_3$ cho.[e]	3.6-3.7 (br)	8.0±1.3	1.8±0.7	1.3×10^{-6}				5.1±1.9	1.1±0.8	1.8×10^{-2}
Glyc.-C1,3H[e]	4.0-4.1 (br)							8.0±2.3	1.8±0.8	6.6×10^{-3}
Glyc.-C1,3H'[e]	4.2-4.3 (br)	3.7±0.9	1.2±0.6	4.5×10^{-4}				7.8±2.3	1.7±0.8	7.1×10^{-3}
Glyc.-C2H[e]	5.1-5.2 (br)				16.4±2.0	2.1±0.7	7.4×10^{-10}	33.9±4.6	3.6±1.1	4.6×10^{-5}
HC=CH[e]	5.2-5.4 (br)	13.8±5.0	0.8±0.6	1.4×10^{-2}	19.2±3.2	1.5±0.6	2.2×10^{-6}	25.5±6.1	1.9±0.8	2.5×10^{-3}
$(CH_2)_n/CH_3$[e]					6.6±1.0	1.6±0.6	2.9×10^{-7}	10.2±2.7	1.8±0.8	5.8×10^{-4}
$HC=CH/CH_3$[e]		8.7±2.9	0.9±0.6	1.9×10^{-3}	10.7±1.9	1.4±0.6	5.2×10^{-6}			
Protein resonances:										
N-acetyl glycoproteins[d]	2.0 (s)				8.9±2.1	1.1±0.6	1.5×10^{-4}	10.7±3.3	1.2±0.8	8.7×10^{-3}
Albumin-lysyl[d]	2.9-3.0 (br)	-6.5±2.1	-1.0±0.6	4.0×10^{-3}	-4.9±1.3	-1.1±0.6	2.9×10^{-4}	-7.2±2.5	-1.3±0.8	1.5×10^{-2}
Proteins NH[f]	5.5-9.0 (br)							-7.0±0.8	-4.1±1.1	2.1×10^{-6}

In the 2nd T, a few changes seem to be partially reverted (namely for alanine, tyrosine, citrate and lactate) and, in the 3rd T, only glutamine (decrease) and creatinine (increase) are seen to change. Conversely, changes in macromolecules become gradually more relevant as pregnancy proceeds (Table 5.1, Figure 5.4), with increases in lipids (viewed by CH_3, $(CH_2)_n$, $=CHC\underline{H}_2CH=$, glyceryl-C1,3H' and HC=CH resonances) and choline metabolites (viewed by $-N(CH_3)_3$ and $-C\underline{H}_2N(CH_3)_3$

resonances), while consistent variations are seen for albumin-lysyl (decrease) and N-acetyl glycoproteins (increase) resonances.

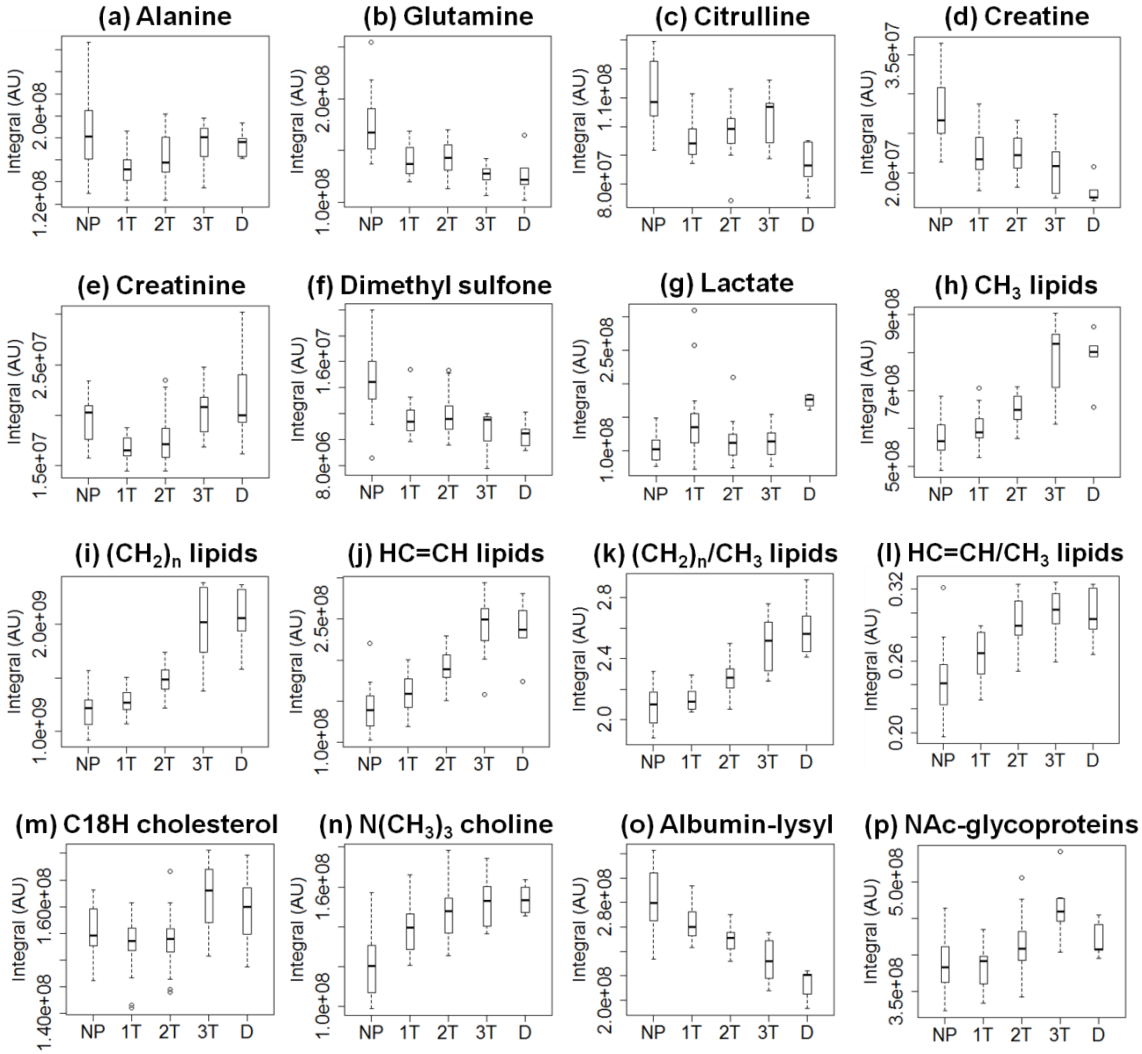


Figure 5.4. Boxplots illustrating the variations (PQN normalized) observed for some metabolites varying from the non-pregnant state and across pregnancy (graphs follow the peak order listed in Table 5.1 i.e. amino acids, other low M_W metabolites, lipids and proteins). The variations represented were measured in the CPMG spectra for all compounds with the exception of lipids and NH proteins, for which the diffusion-edited and standard spectra, respectively, were considered. Graphs corresponding to lipids, choline and cholesterol correspond to the integration of all environments detected for each family i.e. total lipids, total choline-compounds and total cholesterol.

Closer inspection of the diffusion-edited spectra (not shown) provided additional information on the changes affecting lipoproteins and, based on reported assignments (Pinto et al., 2014c) and STOCSY results driven by different methyl resonances (Figure 5.5), distinct lipoprotein spin systems were identified and subsequently monitored separately (Table 5.2).

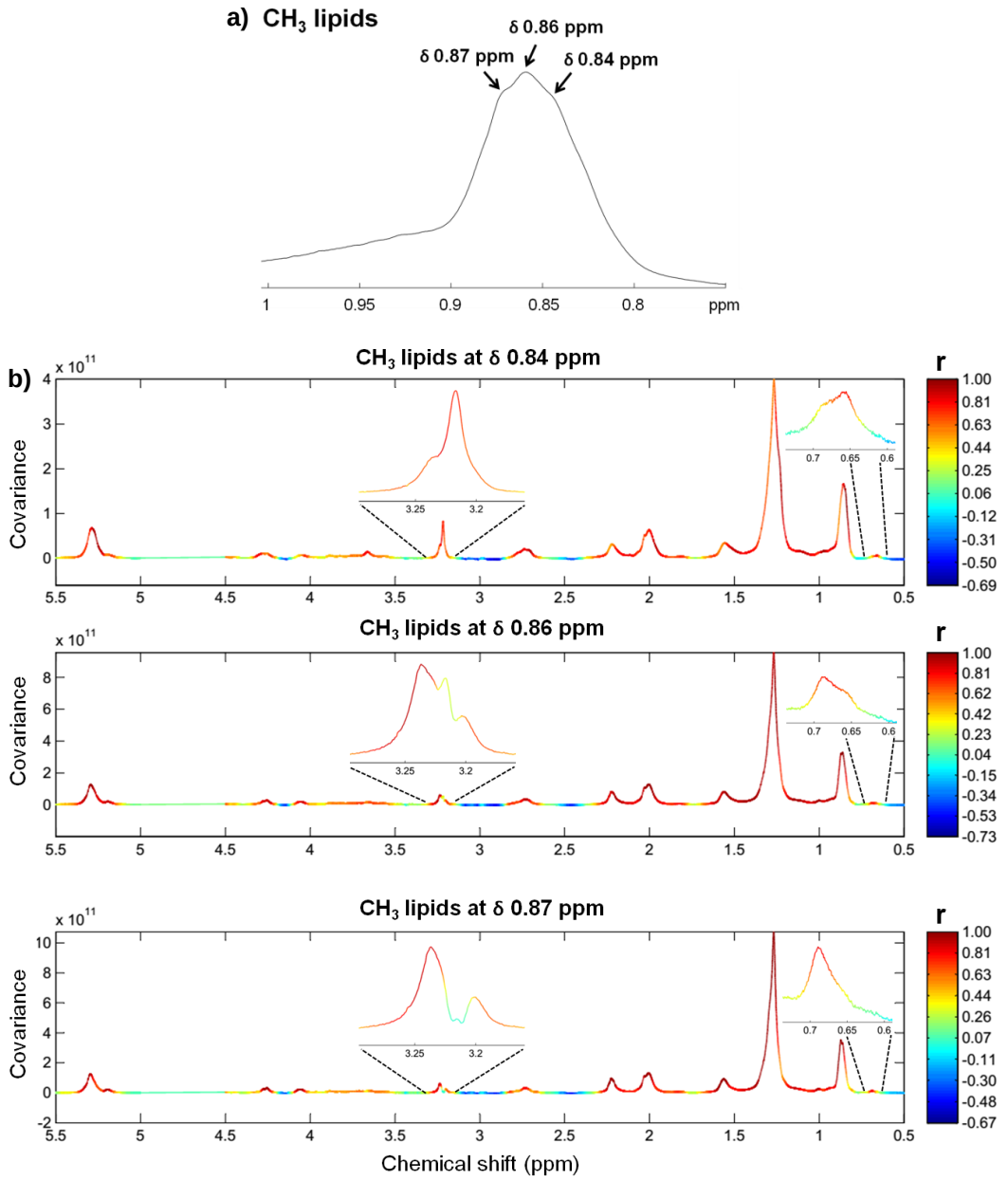


Figure 5.5. a) Average δ 0.75-1.0 ppm region of diffusion-edited ^{1}H NMR spectra of blood plasma of pregnant women. b) 1D STOCSY of diffusion-edited ^{1}H NMR spectra of blood plasma obtained using the CH_3 lipids resonances at 0.84, 0.86 and 0.87 ppm as the root peaks, respectively, with color scale expressing the correlation (r) value.

In this way, distinct C18H cholesterol and -$N(CH_3)_3$ choline resonances were identified as arising either from HDL or from LDL+VLDL (Table 5.2) thus contributing to the respective spin systems: a) for HDL: C18H cholesterol δ 0.65 (broad), CH_3 δ 0.85, $(CH_2)_n$ δ 1.23, other CH_2 groups δ 1.50, 1.98, 2.20, 2.75, -$N(CH_3)_3$ choline δ 3.20 (here named choline A) and δ 3.22 (named choline B), and HC=CH δ 5.27; and b) for LDL+VLDL: C18H cholesterol δ 0.69 (broad), CH_3 δ 0.86 and δ 0.87, $(CH_2)_n$ δ 1.27,

other CH_2 groups δ 1.57, 2.01, 2.22, 2.75, -$N(CH_3)_3$ δ 3.20 (choline A) and δ 3.24 (named choline C), and HC=CH δ 5.30. Choline assignments are partially in agreement with previous reports based on purified lipoprotein subfractions (Lan et al., 2007), which identified choline resonances at δ 3.21 and δ 3.22 respectively as SM and PC in HDL, and choline resonances at δ 3.24 and δ 3.25 for SM and PC in LDL. In this study, it is apparent that, in blood plasma, both HDL and LDL contribute to the δ 3.20 choline resonance (choline A). Free choline (typically resonating at 3.19 ppm (Sitter et al., 2002; Pinto et al., 2014c)) is not identified, as expected, based on its low concentration in the plasma of pregnant women (14-17 μM, i.e. below the detection limit of NMR), compared to that found for phospholipid-bound choline (3300-3500 μM) (Ilcol et al., 2002).

Table 5.2. List of specific variations of HDL and LDL+VLDL, measured in the diffusion-edited ^{1}H NMR spectra of blood plasma of NP state to 1st T, from 1st to 2nd T, and from 2nd to 3rd T. [a] %Variation determined as $(\bar{X}_a - \bar{X}_b)/\bar{X}_b \times 100$, where $\bar{X}$ denotes the average for each group *a* and *b*. [b] E.S.: effect size determined as in (Berben et al., 2012); values in brackets correspond to high uncertainties. [c] Significance level 95% (p-value < 0.05).

Integrated lipoprotein resonances/ppm	NP *vs.* 1st T			1st *vs.* 2nd T			2nd vs. 3rd T		
	%Variat.[a]	E.S.[b]	p-value[c]	%Variat.[a]	E.S.[b]	p-value[c]	%Variat.[a]	E.S.[b]	p-value[c]
HDL									
C18H cholest./δ 0.65							3.6±1.4	1.0±0.7	2.8×10^{-2}
CH_3/δ 0.85	7.7±2.5	0.9±0.6	5.0×10^{-3}	6.3±1.8	0.9±0.6	1.5×10^{-3}	9.4±2.7	1.4±0.8	5.6×10^{-3}
$(CH_2)_n$/δ 1.23	7.7±2.5	0.9±0.6	6.3×10^{-4}	11.8±2.3	1.3±0.6	1.5×10^{-5}	13.1±4.4	1.4±0.8	1.4×10^{-2}
$N(CH_3)_3$ cho. A/δ 3.20	3.6±1.5	0.7±0.6	1.8×10^{-2}				8.2±1.8	1.6±0.8	4.8×10^{-4}
$N(CH_3)_3$ cho. B/δ 3.22	23.8±5.2	1.3±0.6	2.6×10^{-4}	10.1±3.5	0.7±0.5	8.4×10^{-3}			
HC=CH/δ 5.27	16.0±4.9	0.9±0.6	5.4×10^{-3}	16.8 ± 3.0	1.4±0.6	4.4×10^{-6}	18.8±5.2	1.6±0.8	5.2×10^{-3}
$(CH_2)_n/CH_3$ ratio	6.4±1.5	1.3±0.7	2.5×10^{-4}	5.2±1.0	1.4±0.6	4.3×10^{-6}			
$HC=CH/CH_3$ ratio	8.2±2.8	0.9±0.6	8.5×10^{-3}	10.1±1.7	1.5±0.6	9.8×10^{-7}	8.3±2.9	1.2±0.8	1.5×10^{-2}
LDL+VLDL									
C18H cholest./ δ 0.69							8.0±1.9	2.1±0.8	1.7×10^{-3}
CH_3/δ 0.86; 0.87				10.0±2.2	1.2±0.6	1.0×10^{-4}	33.2±6.2	2.8±0.9	6.1×10^{-4}
$(CH_2)_n$/δ 1.27				19.4±3.1	1.5±0.6	8.0×10^{-7}	50.2±9.3	2.7±0.9	8.0×10^{-4}
$N(CH_3)_3$ cho. A/δ 3.20	3.6±1.5	0.7±0.6	1.8×10^{-2}				8.2±1.8	1.6±0.8	4.8×10^{-4}
$N(CH_3)_3$ cho. C/δ 3.24	7.1±2.6	0.8±0.6	1.3×10^{-2}				21.1±4.4	2.2±0.9	9.0×10^{-4}
HC=CH/δ 5.30	12.0±5.4	(0.6±0.6)	4.4×10^{-2}	21.4±3.7	1.4±0.6	3.5×10^{-6}	31.2±7.4	1.9±0.8	2.5×10^{-3}
$(CH_2)_n/CH_3$ ratio				8.5±1.3	1.7±0.6	4.8×10^{-8}	12.1±3.0	1.9±0.8	2.4×10^{-3}
$HC=CH/CH_3$ ratio	8.7±3.1	0.8±0.6	1.2×10^{-2}	10.6±2.1	1.3±0.6	2.2×10^{-5}			
HDL/LDL+VLDL									
CH_3/CH_3 ratio							-16.9±3.5	-2.1±0.9	2.4×10^{-4}

Based on the variations noted for HDL and LDL+VLDL spin systems (Table 5.2 and heatmap in Figure 5.6A), it becomes clear that both lipoprotein levels increase towards the 3rd T, HDL increasing from earlier on in pregnancy but to a lesser extent, compared to LDL+VLDL. The 17% decrease in the methyl resonance ratio CH_3(HDL)/ CH_3(LDL+VLDL) (Figure 5.6B), in the 3rd T, reflects the more marked increase of LDL+VLDL in late pregnancy (Table 5.2).

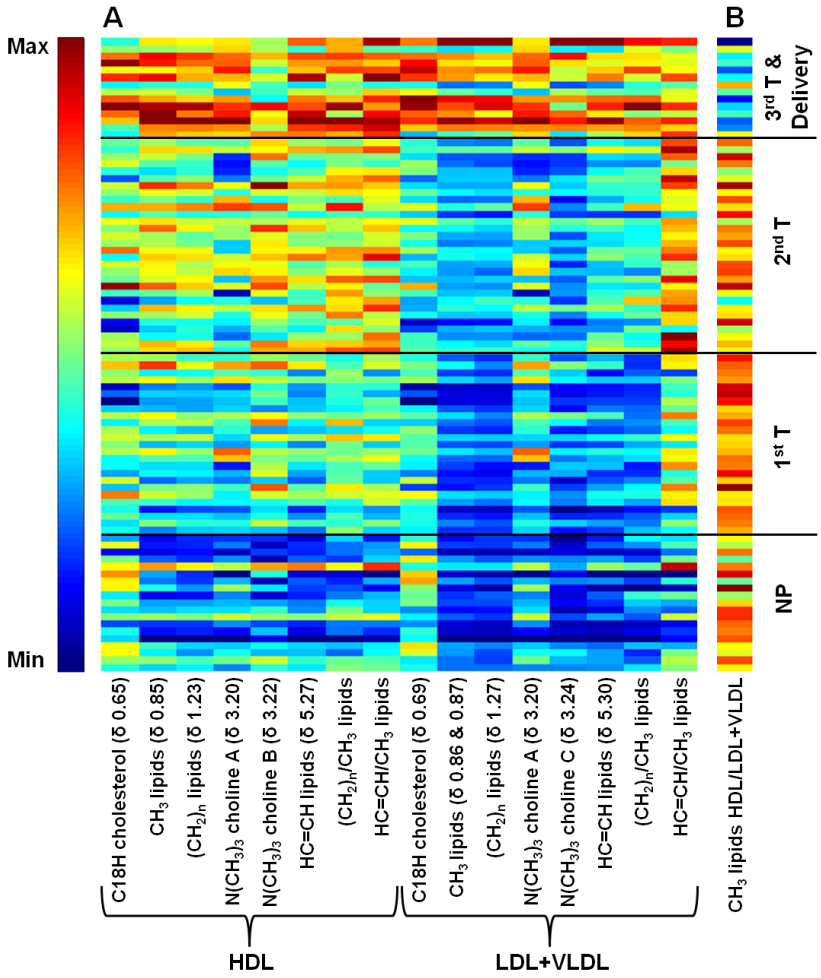


Figure 5.6. Heatmaps of A) normalized integrals of the different lipoprotein resonances (HDL and LDL+VLDL) and corresponding choline and cholesterol environments and B) ratio of methyl resonances in HDL and LDL+VLDL: CH_3(HDL)/CH_3(LDL+VLDL), across pregnancy. Rows represent subjects and columns lipoprotein resonances (or ratios). Integrals are shown using a color scale from minimum (dark blue) to maximum (dark red).

Furthermore, both lipoprotein types change across pregnancy in terms of FA characteristics: increased average chain length and unsaturation degree, respectively given by $(CH_2)_n/CH_3$ and $HC=CH/CH_3$ ratios. Such changes seem to follow different dynamics since 1) average FA chain length increases up to the 2nd T for HDL, and only later in pregnancy for LDL+VLDL and 2) FA unsaturation increases throughout pregnancy for HDL and only in early pregnancy for LDL+VLDL.

5.2. Correlation of plasma changes with urine

Finding metabolite correlations should give additional information on pregnancy metabolism, although the origin of correlations and inter-metabolite distance in metabolic space may not be defined without further biochemical studies. The correlation map in Figure 5.7, panel A shows that, besides correlating positively with

each other, HDL and LDL+VLDL lipoproteins correlate negatively with albumin-lysyl (r -0.80 and -0.72) and total protein (viewed by protein NH resonances, r -0.73 and -0.83), suggesting an inverse metabolic relationship between the two plasma macromolecule classes.

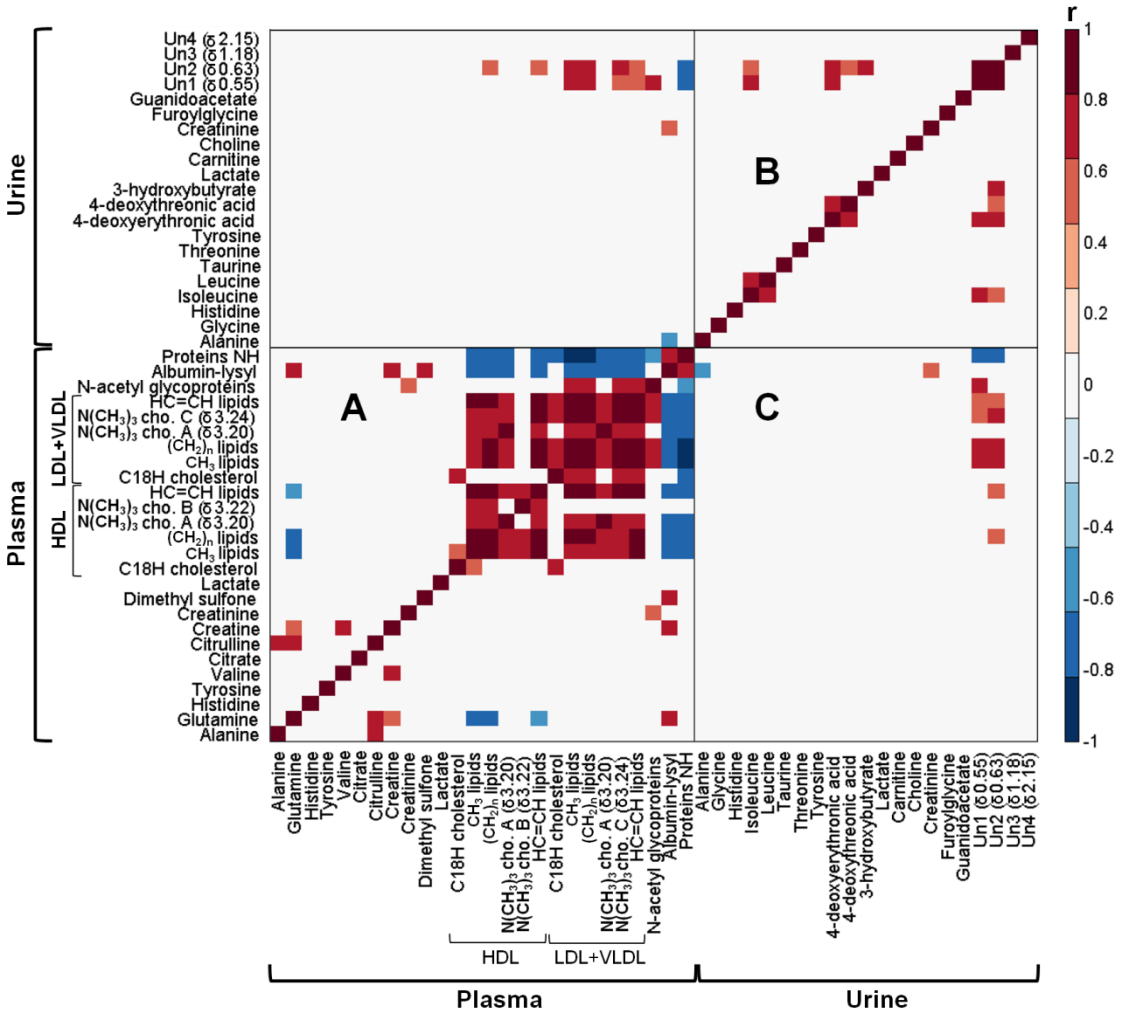


Figure 5.7. Correlation analysis of metabolite integrals that vary significantly in maternal blood plasma and urine throughout pregnancy. Panels A, B and C show metabolite correlations in plasma/plasma, in urine/urine and in plasma/urine respectively. Only metabolite correlations with $|r| \geq 0.6$ and $p < 0.05$ were considered. These correlations were confirmed by reviewing the scatter plots of each pair of correlated integrals.

Inter-metabolite correlations in plasma are comprehensively illustrated in Figure 5.8 (left), where line thickness and colour reflect correlation strength and sign respectively. The negative correlations between lipoproteins and albumin/total protein form an evident correlation pattern, along with strong (positive) correlations of HDL with LDL+VLDL (r 0.83) and of albumin with NH resonances (r 0.75), the latter observation being consistent with albumin being the most abundant protein in plasma. Interestingly,

LDL+VLDL seems specifically (and positively, r 0.71) correlated to N-acetyl-glycoproteins (and these, in turn, to creatinine with r 0.60). Albumin levels are apparently also related to changes in creatine (r 0.66) (in turn correlated to valine (r 0.66) and glutamine (r 0.60)), glutamine (r 0.71) (in turn correlated to alanine (r 0.60) and citrulline (r 0.64)) and dimethyl sulfone (r 0.70).

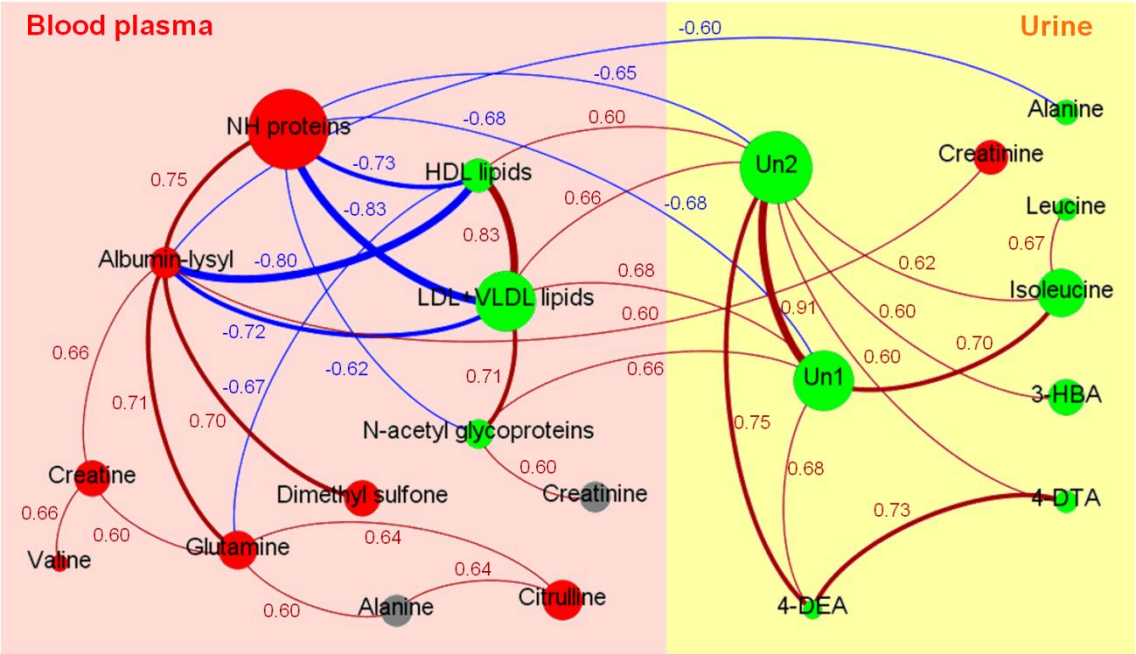


Figure 5.8. Correlation network of maternal plasma and urine metabolites observed to change throughout pregnancy, based on Pearson's correlation coefficients ($|r| \geq 0.6$ and $p < 0.05$). Node size denotes the highest effect size across pregnancy for each metabolite. Node colours indicate direction of variation throughout pregnancy with red for decrease, green for increase and gray for metabolites that showed both decrease and increase. Dark red lines correspond to positive correlations, whereas blue lines correspond to negative correlations. Line thickness reflects the magnitude of the correlation coefficients ($|r| \geq 0.6$, thin; $0.7 \leq |r| < 0.8$, medium; $|r| \geq 0.8$, broad).

Metabolite correlations between maternal plasma and urine and within urine were also calculated (Figure 5.7 and Figure 5.8). The strongest correlations within maternal urine during pregnancy (Figure 5.7, panel B and Figure 5.8) confirmed previous reports (Diaz et al., 2013a) and the importance of the pregnancy urinary markers 4-deoxyerythronic acid (4-DEA), 4-deoxythreonic acid (4-DTA) and unassigned spin systems Un 1 (δ 0.55, s) and Un 2 (δ 0.63, s). These unassigned systems have been suggested to arise from excreted bile acids (Diaz et al., 2013a) but confirmation is still lacking. Metabolic relationships between excreted and circulatory metabolomes, during pregnancy, are illustrated here for the first time to our knowledge through plasma/urine correlations (Figure 5.7, panel C and Figure 5.8). The fact that plasma and urine

samples correspond to partially overlapped cohorts means that measured correlations reflect average cohort behaviour, rather than individual behaviour. Lipid metabolism seems to be related to the excretion of Un 1 (correlated to LDL+VLDL alone with r 0.68) and Un 2 (correlated to both HDL and LDL+VLDL, with r 0.60 and 0.66). Furthermore, LDL+VLDL metabolism seems to also specifically relate to glycoproteins (r 0.71). Both Un 1 (r -0.68) and Un 2 (r -0.65) are themselves negatively correlated to changes in total plasma proteins (viewed by the NH spectral region), but not to albumin (for which only weak correlations are seen, with r 0.50 and r 0.59), probably reflecting a relationship between the unassigned compounds and other proteins or peptide hormones known to change across pregnancy (e.g. human Chorionic Gonadotropin (hCG), known to decrease in the 3rd T (Edelstam et al., 2007)). In turn, albumin shows apparent positive and negative correlations to excreted creatinine (r 0.60) and alanine (r -0.60), respectively (Figure 5.7 and Figure 5.8).

5.3. Proposed metabolic interpretation of plasma changes

Low M_w compounds

The decreased levels of circulating amino acids during pregnancy is well documented and explained in terms of the increased placental transfer of amino acids to the fetus, favouring nitrogen conservation for fetal growth (King, 2000; Herrera and Ortega, 2008). The lower levels noted here for valine, alanine, glutamine, tyrosine, histidine and citrulline in the 1st T identify these amino acids as particularly important in serving the fetus, early on in pregnancy. Previous reports have also associated decreases in alanine and citrulline to the slowing down of urea synthesis late in pregnancy (King, 2000). Indeed, an association between alanine and citrulline is apparent in this work, possibly involving glutamine too (Figure 5.8). In the 2nd T, lower amounts of alanine and tyrosine seem to be needed, since maternal circulatory levels are partially recovered. Other amino acids are generally maintained at a low level until the end of pregnancy, an enhanced use of glutamine being noted in the 3rd T. Our correlation results suggest a strong positive correlation linking glutamine to albumin metabolism, the exact nature of which remains unknown at this stage. The amino acid variations noted here are in broad agreement with the literature, although the exact amino acid fluctuations accompanying healthy pregnancies seems to vary significantly between studies (Di Giulio et al., 2004; Dasarathy et al., 2010), a fact that may be explained by either different analytical aspects (e.g. experimental/analytical design, sample numbers) and/or different population phenotypes. Other small metabolites that

vary include citrate, dimethyl sulfone, lactate, creatine and creatinine. To our knowledge, no changes in circulating citrate and lactate have been reported in connection to pregnancy, these metabolites probably indicating a shift in maternal energy metabolism in the 1st T, partially compensated in the 2nd T. The early decrease in dimethyl sulfone, also newly observed here, may be indicative of an early change in intestinal bacterial metabolism and/or endogenous methanethiol (a known source of dimethyl sulfone in blood) metabolism (Dasarathy et al., 2010), although a dietary source cannot be entirely ruled out. Regarding creatine and creatinine, only alterations in the latter have been reported to our knowledge in a pregnancy context, namely lower circulating levels and increased excretion, together with ammonia and uric acid (King, 2000; Creasy et al., 2009; Cheung and Lafayette, 2013). A rise in glomerular filtration rate (GFR), without substantial alterations in the production of creatinine and urea, has been advanced to explain these variations. In this work, both plasma creatine and creatinine decrease in the 1st T and this seems to be compensated for creatinine later in pregnancy, possibly due to a GFR decrease known to occur near term (Creasy et al., 2009). The absence of a visible correlation of creatine with creatinine (either in plasma or between plasma and urine) suggests that a dynamically complex network involves these two metabolites, an issue to be investigated in future targeted studies.

High M_w compounds

Our results confirm the expected increases in lipoproteins and lipoprotein-cholesterol, described in several reports (Butte, 2000; Lippi et al., 2007; Herrera and Ortega, 2008; Hadden and McLaughlin, 2009; Mankuta et al., 2010). Maternal fat deposition in early and mid-pregnancy is used as an energy source for the mother so that glucose is spared for the fetus (Hadden and McLaughlin, 2009). In the 3rd T, lipolysis is promoted and lipogenesis suppressed, leading to increased levels of triacyglycerols (in both HDL and LDL) and VLDL particles in maternal blood plasma (Herrera and Ortega, 2008; Hadden and McLaughlin, 2009). Evidence of this metabolic switch is clearly illustrated by the high 3rd T levels of HDL and LDL+VLDL seen here, although the corresponding dynamic evolution is quite distinct, leading to a 17% decrease in HDL/LDL+VLDL in the 3rd T. Our results suggest that the onset of lipolysis late in pregnancy seems to preferentially produce longer and more unsaturated FAs, which become components of 3rd T HDL and LDL+VLDL.

Furthermore, our correlation studies have shown that the increases in HDL and LDL+VLDL display a strong correlation with the observed albumin decrease across

pregnancy (the latter variation having been reported previously (Whittaker and Lind, 1993; Bacq et al., 1996; King, 2000)) and, interestingly, to the excretion of the still unassigned pregnancy marker Un 2 (δ 0.63 ppm, s). On the other hand, LDL+VLDL metabolism seems to have its own specific signature (distinct from HDL), involving a link to glycoprotein metabolism and excretion of another unassigned pregnancy marker Un 1 (δ 0.55 ppm, s). These observations reveal specific links between different lipoproteins and other protein classes in plasma, and between circulating and excreted metabolomes. The increase in N-acetyl-glycoproteins may reflect a number of proteins (α1-acid glycoprotein, α1-antitrypsin and haptoglobulin (Torri et al., 1999)), previously reported as changing during pregnancy (Honda et al., 1990; Larsson et al., 2008). The decrease in albumin seen here has been reported in several studies (Whittaker and Lind, 1993; Bacq et al., 1996; King, 2000), although different reasons have been advanced for this observation. The most common explanation relates to an increase in plasma volume during pregnancy (Bacq et al., 1996), however, in this work, data normalization should account for such a general concentration effect. Indeed, other possible origins for plasma albumin depletion in pregnancy may relate to reported changes in albumin metabolism (Whittaker and Lind, 1993), increased excretion (Higby et al., 1994) or increased transfer to the fetus (Sutcliffe, 1975). The origin of the apparent correlation of albumin with plasma creatine, glutamine and dimethyl sulfone and with excreted alanine and creatinine is unclear at this stage, although a possible relation to the selective action of this protein as a transport agent in blood circulation may be advanced.

5.4. Conclusions

This work describes the first NMR metabolomic study of pregnancy using maternal plasma and correlation between plasma and urine metabolites. The expected decrease in circulating amino acids early in pregnancy was confirmed with six amino acids identified as being required to larger extents by the fetus. Novel observations include early changes in citrate, lactate and dimethyl sulfone levels, possibly due to an metabolic energy shift and gut microflora adjustments taking place in the 1st T. Alteration in creatine levels was also noted, along with creatinine changes with the latter known to reflect changes in GFR during pregnancy.

Plasma HDL and LDL+VLDL levels were confirmed to increase throughout pregnancy, but at different rates and accompanied by increases in fatty acid chain length and unsaturation degree. Distinct correlation patterns suggested a) an inverse

link between HDL and LDL+VLDL and plasma albumin, along with a direct link with excreted pregnancy markers Un 1 (singlet at δ 0.55) and Un 2 (singlet at δ 0.63); b) a direct specific relationship between LDL+VLDL and N-acetyl-glycoproteins, together with excreted Un 1 and c) a number of correlations of plasma albumin with specific circulating (possibly due to selective transportation) and excreted small metabolites. These results unveil specific lipoprotein/protein metabolic aspects of pregnancy, with an impact on the excreted metabolome, and as such provide an interesting lead for the further understanding of pregnancy metabolism.

Chapter 6. First and second trimester maternal plasma for diagnosis of chromosomal disorders and trisomy 21

The results presented in this chapter were fully submitted/published in the articles below and are here adapted slightly onto the form of a thesis chapter, for the sake of clarity:

Pinto, J.; Almeida, L. M.; Martins, A. S.; Duarte, D.; Domingues, M. R. M.; Barros, A. S.; Galhano, E.; Pita, C.; Almeida, M. C.; Carreira, I. M.; Gil, A. M. Fetal chromosomal disorders impact on maternal metabolome: towards new biomarkers? *American Journal of Obstectrics & Gynecology* **2015**, under revision.

Pinto, J.; Maciel, E.; Melo, T. S.; Domingues, M. R. M.; Galhano, E.; Pita, C.; Almeida, M. do C.; Carreira, I. M.; Gil, A. M. Maternal Plasma Phospholipids Are Altered in Trisomy 21 Cases and prior to Preeclampsia and Preterm Outcomes. *Rapid Communications in Mass Spectrometry* **2014**, *28*, 1635–1638.

6.1. Analysis of plasma and total lipid extracts by NMR spectroscopy

Abstract

Objectives: This study aimed at determining the impact of fetal CD, including T21, on 1st and 2nd T maternal blood plasma, to identify 1) time course metabolic adaptations to the conditions and 2) possible new plasma biomarkers. Furthermore, a definition of a joint circulatory (plasma) and excretory (urine) metabolic description of 2nd T CD was sought. Study Design: Blood plasma was obtained for 119 pregnant women: 74 controls and 45 CD cases, including 22 T21 cases. Plasma and lipid extracts were analysed by NMR spectroscopy and data was handled by variable selection and MVA. Correlation analysis was used on a concatenated plasma/urine matrix descriptive of 2nd T CD, based on previously obtained urine data. Results: CD cases were accompanied by enhanced lipid β-oxidation (increased ketone bodies) and underuse of glucose, pyruvate and citrate. Lower circulating HDL levels were noted, along with changes in i) proline and methanol in the 1st T, and ii) urea, creatinine and LDL+VLDL in the 2nd T. T21 cases were indistinguishable from other CD cases in the 1st T, while in the 2nd T increased methanol and albumin may be T21-specific. Furthermore, 1st T lipid extracts of T21 showed decreased levels of 18:2 fatty acids, whereas in the 2nd T lower levels of 20:4 and 22:6 fatty acids were noted, possible indicative of inflammation mechanisms. In both trimesters, high classification rates for CD (88-89%) and T21 (85-92%) generally relied on variable selection of NMR data. Plasma/urine correlations confirmed most metabolic deviations and unveiled possible new ones regarding LDL+VLDL, sugar and gut-microflora metabolisms. Conclusions: This work partially confirmed previously reported data on 1st T T21 and provided additional information on time course metabolic changes accompanying CD and T21, in particular regarding plasma lipid composition. These results demonstrate the potential of plasma metabolomics in monitoring and characterizing CD cases in the clinic, however, validation in larger cohorts and demonstration of putative biochemical hypothesis are required before applications may be envisaged.

Brief state of the art

Current diagnostic methods for CD comprise invasive procedures such as CVS and amniocentesis performed at 10-13 g.w. and 15-20 g.w., respectively. These procedures involve some risk for mother and fetus (0.5-1% fetal loss rate (Tabor and Alfirevic, 2010)) and, currently, non-invasive screening tests are used in the 1st and 2nd T based on ultrasound and maternal serum biomarkers (β-hCG, PAPP-A, AFP, uE3 or inhibin

A) (Nicolaides, 2011). 1st T screening combines maternal age, fetal nuchal translucency thickness and serum free β-hCG and PAPP-A, with detection and false positive rates (FPR) of 85-95% and 5% for T21, and > 90% and 1% for other aneuploidies (T18 and T13). In the 2nd T, biomarker combination provides 70-75% detection rates and 5% FPR for T21 (Nicolaides, 2011). Analysis of cffDNA in maternal blood has a higher performance, with 1st T detection rates of 98.6-100%, 100% and 91.7% for T21, T18 and T13, respectively, and FPR < 1% (Walsh and Goldberg, 2013). However, the high cost of such tests and the relative low detection rates for T13 presently hinder their use as replacements of either CVS or amniocentesis. Hence, there is still a need for biomarkers with improved reliability, as well as for a deeper knowledge of how CD impact on pregnancy metabolism, thus enlightening the conditions´ mechanisms and interaction between genetic and metabolic traits.

The metabolic impact of CD on 2nd T amniotic fluid was low, viewed by NMR spectroscopy (Graça et al., 2010), but a subsequent HPLC study unveiled amino acid changes in T21 cases (Amorini et al., 2012). NMR of 2nd T maternal urine resulted in CD classification with 87% sensitivity and 84% specificity, suggesting higher glucose excretion and amino acid alterations (Diaz and Pinto et al., 2011). T21 cases were distinguished from other CD by possible changes in energy, nucleotide and amino acid metabolisms. A subsequent MS study of 1st and 2nd T maternal urine of T21 revealed increased dihydrouracil (possibly due to liver abnormalities) and decreased progesterone (probably reflecting impairment of placental function due to oxidative stress) (Trivedi and Iles, 2014b). Regarding maternal blood, an initial indication of lipid changes in 2nd T maternal plasma of T21 (Diaz and Pinto et al., 2011) was later confirmed by lipidomics, which revealed changes in LPC and SM, along with ether-linked PC and PE species related to lipid oxidation (Pinto et al., 2014d). NMR of T21 1st T maternal blood reported changes in 11 metabolites, particularly in 3-hydroxybutyrate, 3-hydroxyisovalerate and 2-hydroxybutyrate (75% sensitivity, 86.2% specificity) (Bahado-Singh et al., 2013b). These results were interpreted as reflecting oxidative stress, poor myelination and neurotoxicity of the brain in T21 fetuses. A similar study of T18 advanced a combination of 2-hydroxybutyrate, glycerol and maternal age for T18 detection (73% sensitivity, 97% specificity), while trimethylamine, threonine and creatine levels seemed to differentiate T18 from T21 (Bahado-Singh et al., 2013c).

In this work, a NMR metabolomics study of 1st and 2nd T maternal plasma of a general CD group and a T21 subgroup is presented in comparison to controls, in order to 1) establish the time course adaptations of maternal/fetal metabolism to the conditions and 2) unveil dynamic metabolic biomarkers. Furthermore, blood plasma and lipid extracts are compared in terms of T21 classification, the latter adding specific

information on lipid composition. Finally, the CD-related metabolic changes measured in 2nd T maternal plasma are correlated to the corresponding urine samples, thus providing a more complete metabolic description of CD in mid-pregnancy.

6.1.1. Metabolic fingerprinting of CD in the 1st trimester

The average CPMG ^{1}H NMR spectrum of control 1st T maternal plasma (Figure 6.1a) is dominated by broad lipoprotein and glucose peaks whereas the corresponding lipid extract spectrum (Figure 6.1b) shows additional information on lipid composition, namely, several different methylene and methine environments, SM and PC choline resonances, TG and PL glyceryl resonances and 18:2, 20:4 and 22:6 FA. PCA of the CPMG ^{1}H NMR spectra of 1st T blood plasma of controls and CD samples (not shown) revealed no trends and the corresponding PLS-DA model had poor performance (Q^2 0.300) (Table 6.1, Figure 6.2a). Variable selection was applied to the spectra and the subsequent PLS-DA 1st T plasma model (Figure 6.3a) was indeed much improved (Q^2 0.700, Table 6.1, Figure 6.2b), the set of variables selected being shown as grey dots in Figure 6.1a.

Table 6.1. MCCV parameters obtained when considering the original (full resolution) ^{1}H NMR spectra and the NMR spectra after variables selection through the intersection of VIP>1, $VIP/VIP_{cvSE} > 1$ and $b/b_{cvSE} > 1$. T: trimester; LV: no. of latent variables; Q^2: most frequent value obtained by MCCV; CR: classification rate, Sens: sensitivity and Spec: specificity; %VS: percentage of variables selected from the original spectra; values in bold refer to robust PLS-DA models.

T	Subjects group	Original spectra LV	Q^2	CR	Sens	Spec	Spectra after variable selection LV	Q^2	CR	Sens	Spec	% VS
1st	***Whole blood plasma***											
	Controls (n=24) /CD (n=15)	2	0.3	70%	41%	89%	1	**0.7**	**88%**	**78%**	**94%**	**34%**
	Controls (n=24) /T21 (n=8)	1	0.2	76%	18%	96%	1	**0.7**	**92%**	**72%**	**99%**	**32%**
	Plasma lipid extracts											
	Controls (n=15) /T21 (n=7)	-	-	-	-	-	1	**0.8**	**87%**	**78%**	**91%**	**25%**
2nd	***Whole blood plasma***											
	Controls (n=49) /CD (n=30)	-	-	-	-	-	2	**0.7**	**89%**	**79%**	**94%**	**31%**
	Controls (n=49) /T21 (n=14)	1	0.3	78%	26%	93%	1	**0.7**	**85%**	44%	**97%**	**32%**
	Plasma lipid extracts											
	Controls (n=15) /T21 (n=13)	2	**0.7**	**94%**	**94%**	**94%**	2	**0.9**	**98%**	**100%**	**97%**	**33%**
	Whole plasma and urine											
	Controls (n=22) /CD (n=26)	1	0.3	63%	62%	64%	1	**0.6**	**90%**	**85%**	**95%**	**32%**
	Controls (n=22) /T21 (n=12)	-	-	-	-	-	1	**0.8**	**95%**	**85%**	**100%**	**32%**

The fact that some selected variables fall in regions of weak/broad resonances or slight baseline nuances (e.g. in the aromatic region) indicates the importance of spectral detail for sample classification. Computed metabolite changes (Table 6.2) showed that CD samples carry lower amounts of HDL lipoproteins, proline and methanol, and increased levels of β-hydroxybutyrate and pyruvate (as well as qualitative increases in acetate, acetone, citrate, creatine and glucose), as clearly seen in a Volcano plot format (Figure 6.3a right).

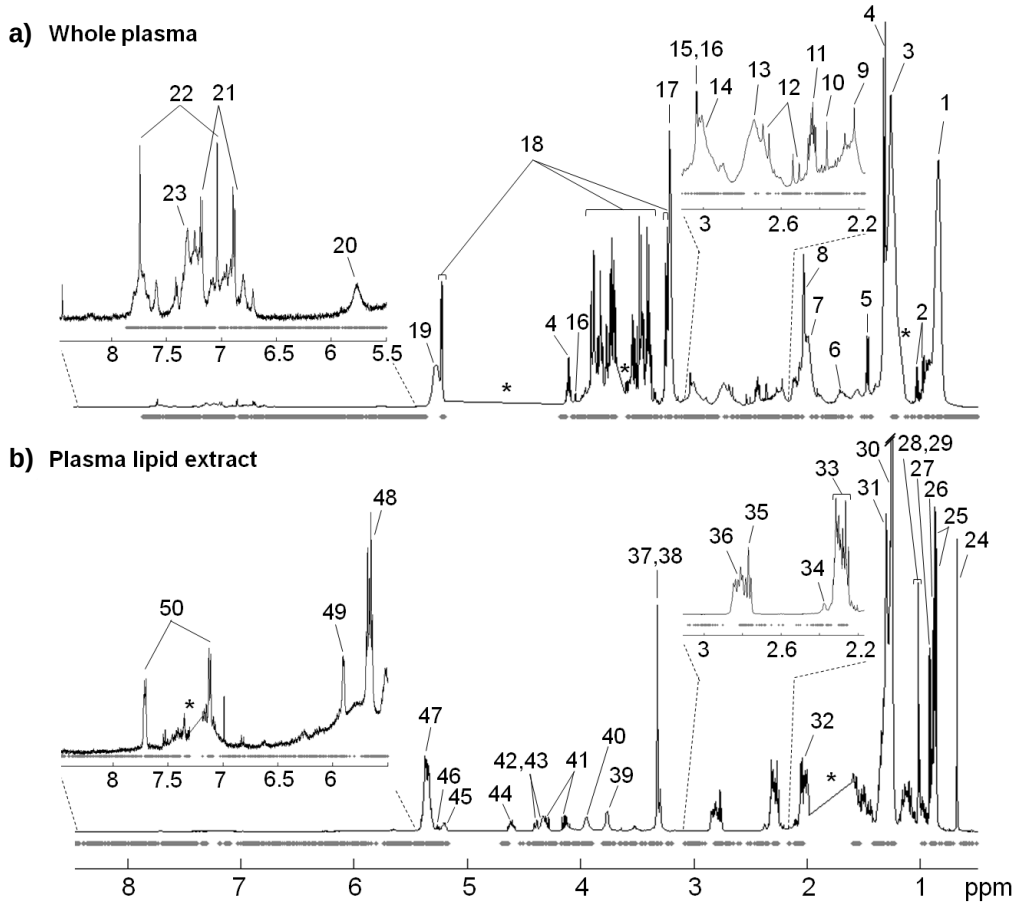
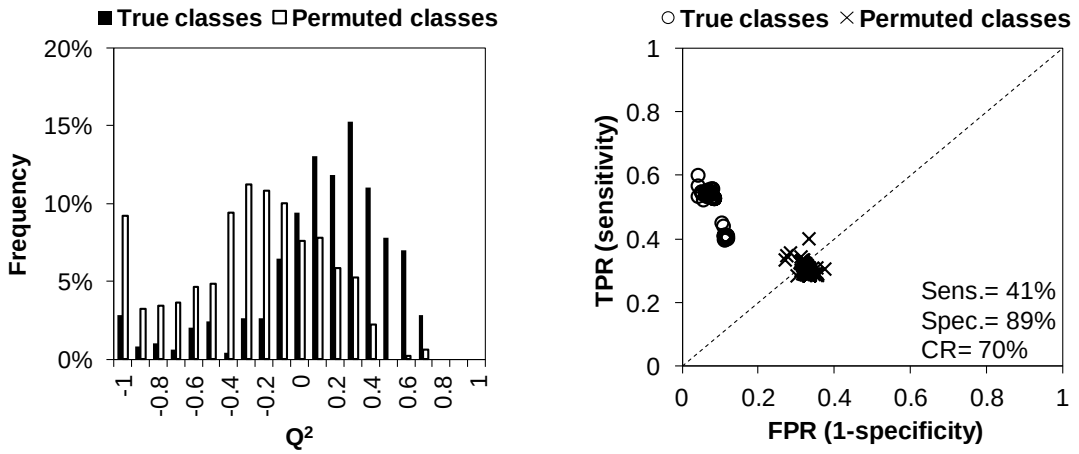


Figure 6.1. a) Average (n=24) 500 MHz CPMG ^{1}H NMR spectrum of 1st T plasma of healthy pregnant women (controls), with indication of the spectral variables (data points) selected in the PLS-DA model controls vs. CD (grey dots under spectrum); b) average (n=15) 500 MHz ^{1}H NMR spectrum of 1st T plasma lipid extracts of controls, with indication of the spectral variables selected in the PLS-DA model of controls vs. T21 (grey dots under spectrum). Assignment: 1- CH_3 lipids, 2- valine, 3- $(CH_2)_n$ lipids, 4- lactate, 5- alanine, 6- $C\underline{H_2}CH_2CH=$ lipids, 7- $C\underline{H_2}CH=$ lipids, 8- N-acetyl glycoproteins, 9- acetone, 10- pyruvate, 11- glutamine, 12- citrate, 13- $=CHC\underline{H_2}CH=$ lipids, 14- albumin-lysyl, 15- creatine, 16- creatinine, 17- $N(CH_3)_3$ choline PL, 18- glucose, 19- HC=CH lipids, 20- urea, 21- tyrosine, 22- histidine, 23- unknown 1 (δ 7.19-7.38, br), 24- $C18H_3$ FC and EC, 25- $C26H_3$ and $C27H_3$ FC and EC, 26- CH_3 FA, 27- $C21H_3$ FC and EC, 28- $C19H_3$ FC, 29- $C19H_3$ EC, 30- $(CH_2)_n$ FA, 31- $=CHCH_2(C\underline{H_2})_n$ FA, 32- $C\underline{H_2}CH=$ FA, 33- CH_2CO FA, 34- CH_2CO 22:6 FA, 35- $=CHC\underline{H_2}CH=$ 18:2 FA, 36- $=CHC\underline{H_2}CH=$ 20:4 FA, 37- $N(CH_3)_3$ choline of SM, 38- $N(CH_3)_3$ choline of PC, 39- $C\underline{H_2}N(CH_3)_3$ PC and SM, 40- glyceryl-$C3H_2$ PL, 41- glyceryl-C1,3H and -C1,3H' TG, 42- glyceryl-$C1H_2$ PL, 43- CH_2-PO PL, 44- C3H EC, 45- glyceryl-C2H PL, 46- glyceryl-C2H TG, 47- HC=CH FA, 48- $=C\underline{H}$-CHOH SM, 49- Unknown 2 (δ 5.91, br), 50- Unknown 3 (δ 7.12, d and δ 7.71, m). *: cut off spectral regions corresponding to water and ethanol peaks in a) and chloroform and water peaks in b).

When comparing T21 cases with controls, the PLS-DA model (after variable selection) and metabolite changes noted were very similar to those describing the general CD group (Figure 6.3b, Table 6.1 and Table 6.2). Since T21 (n=8) cases predominate within the CD group and are indistinguishable from other CD (n=7) in multivariate space (results not shown), no specific T21 signature may be advanced at this stage. Previous reports on 1st T deproteinized sera of T21 and T18 cases (Bahado-Singh et al., 2013b, 2013c) also identified increases in β-hydroxybutyrate, pyruvate and acetone, in addition to other lower abundance metabolites probably made visible by the filtration/extraction procedure employed. Here, direct plasma analysis unveils additional lipid compositional changes. Considering lipid extract spectra, a distinct subset of variables was selected (Figure 6.1b), 1st T extracts being shown to have decreased levels of 18:2 FA and PC (Figure 6.3c).

a) Controls 1st T (n=24) vs. CD 1st T (n=15) – CPMG original spectra



b) Controls 1st T (n=24) vs. CD 1st T (n=15) 1T – CPMG spectra after variable selection

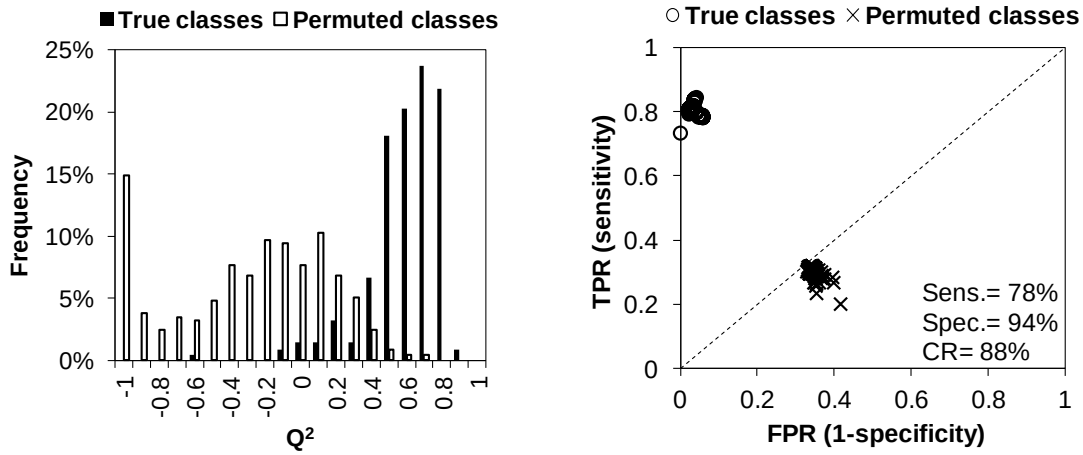


Figure 6.2. Q^2 distributions (left) and ROC plots (right) of true and permuted models obtained for a) the original CPMG ^{1}H NMR spectra of whole blood plasma of controls 1st T (n=24) vs. CD 1st T (n=15) and b) corresponding variable selected spectra.

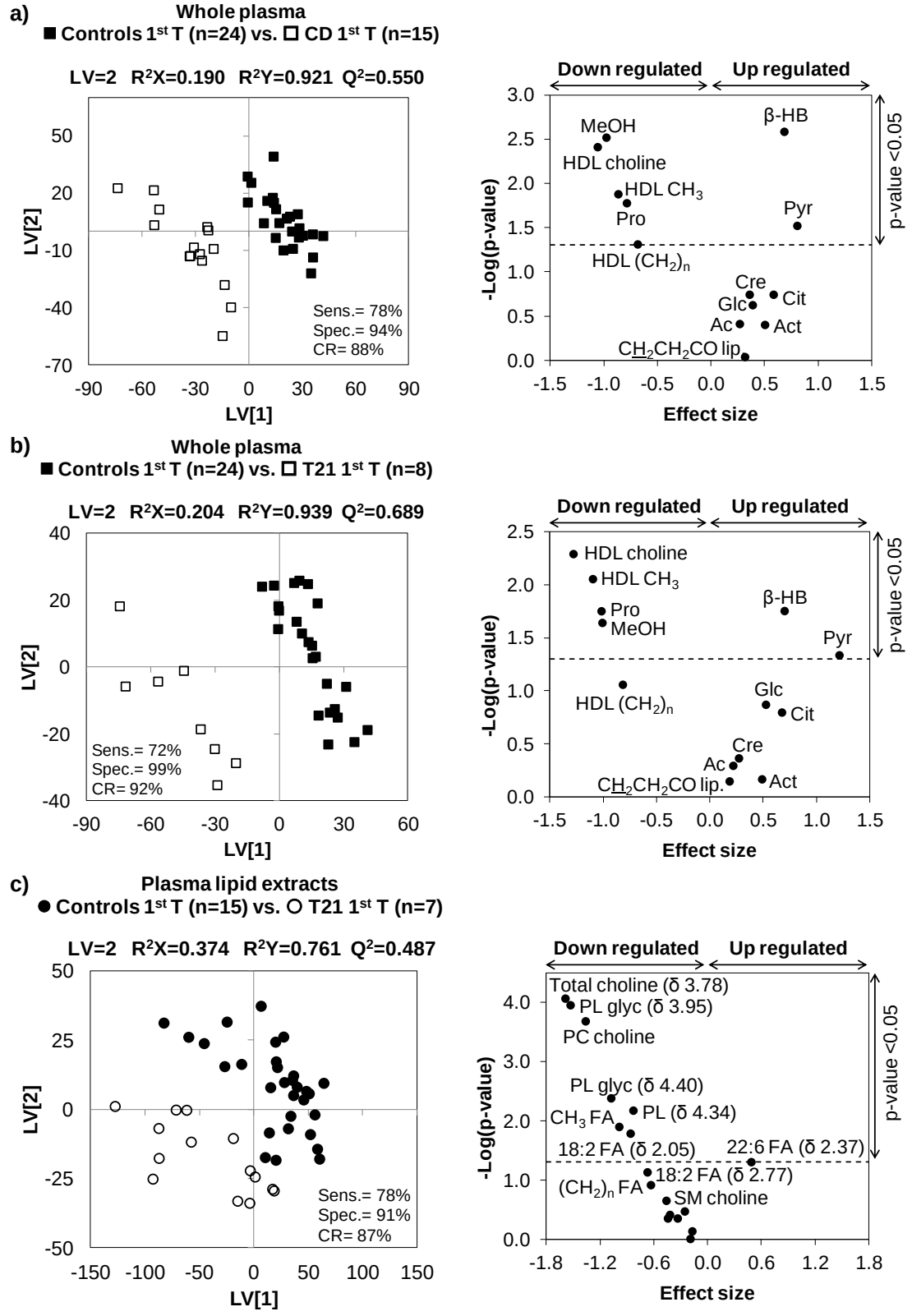


Figure 6.3. PLS-DA scores scatter plots for CPMG ^{1}H NMR plasma spectra (after variables selection) and Volcano plots (effect size vs. –Log (p-value)) representing contributing metabolites/resonances for whole blood plasma of controls 1st T (n=24) vs. a) CD 1st T (n=15) and b) T21 1st T (n=8); c) plasma lipid extracts of controls 1st T (n=15) vs. T21 1st T (n=7). Ac, acetate; Act, acetone; β-HB, β-hydroxybutyrate; Cit, citrate; Cre, creatine; Glc, glucose; glyc, glyceryl; MeOH, methanol; Pro, proline; Pyr, pyruvate. Values in brackets correspond to compound chemical shifts.

6.1.2. Metabolic fingerprinting of CD in the 2nd trimester

A similar strategy was employed for comparing 2nd T CD (and T21) cases and controls. PLS-DA of variable selected data (Figure 6.4a) and subsequent signal integration enabled the identification of a different metabolic profile describing 2nd T CD cases (Table 6.2, Figure 6.4a): decreased urea and acetate, and increased creatinine and citrate. Additional qualitative differences comprised increasing tendencies for LDL+VLDL and alanine, and the absence of changes in creatine and methanol (noted in the 1st T). It is important to note that the exact composition of 1st and 2nd T CD groups is somewhat different (see Table 2.4, page 66), both groups being dominated by T21 (53% and 47% of all cases in 1st and 2nd T respectively) and including small numbers of other disorders; therefore, a possible small contribution from under-represented disorders to the apparent metabolic adaptions/evolution based on average CD profiles is possible.

Regarding the 2nd T T21 subgroup, a large similarity to the whole CD group is again noted, with significant decreased urea, acetate and HDL (Figure 6.4b). However, an increase in glucose becomes clearer (although with high uncertainty, Table 6.2), along with apparently T21-specific changes († in Table 6.2): increased lipid $C\underline{H_2}CH_2CO$ resonance, albumin, methanol and unknown 1 (δ 7.19-7.38), and absence of changes in β-hydroxybutyrate, proline and creatinine. The spectra of 2nd T extracts produced robust PLS-DA models, with or without variable selection (Table 6.2, Figure 6.4c) and identified FA methylenes, PC, total choline and PL glyceryl groups as decreased in T21, similarly to 1st T plasma. However, some differences are observed, namely the now non-significant change in 18:2, decreases in 22:6 and 20:4 FAs, absence of change in SM and increase in unknown 3 (δ 7.71) (Table 6.2, Figure 6.4c). The observation of different metabolic profiles for 2nd T CD and T21 groups, compared to the 1st T, suggests that a dynamic metabolic adaptation to CD occurs across pregnancy, rather than a simple enhancement of the metabolic impact registered in early pregnancy (although a possible contribution of different CD group compositions may not be completed ruled out).

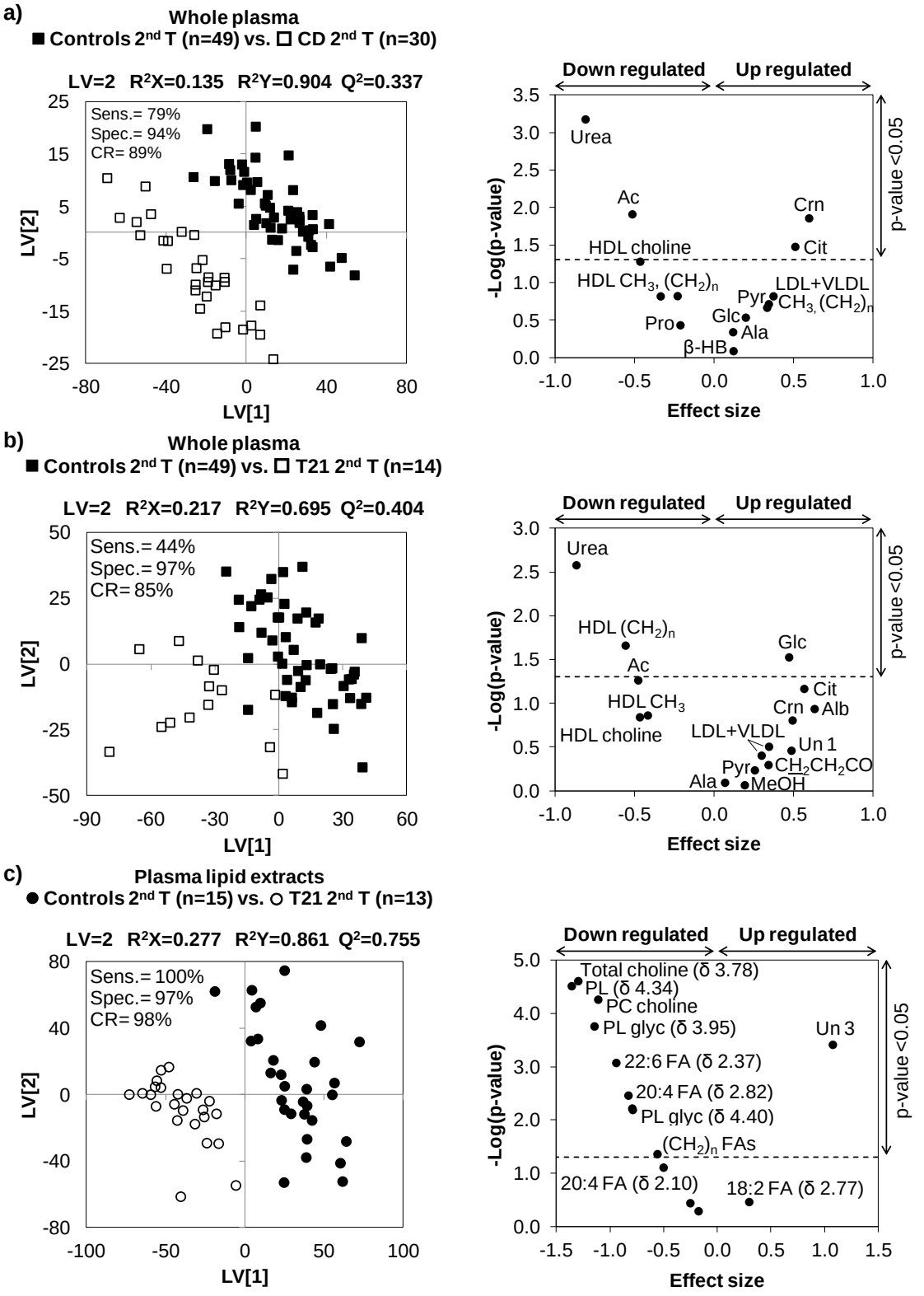


Figure 6.4. PLS-DA scores scatter plots for CPMG ^{1}H NMR plasma spectra (after variables selection) and Volcano plots (effect size vs. –Log (p-value)) representing contributing metabolites/resonances for whole blood plasma of controls 2nd T (n=49) vs. a) CD 2nd T (n=30) and b) vs. T21 (n=14); c) plasma lipid extracts of controls 2nd T (n=15) vs. T21 2nd T (n=13). Ac, acetate; Ala, alanine; Alb, albumin-lysyl; β-HB, β-hydroxybutyrate; Cit, citrate; Crn, creatinine; FAs, fatty acids; Glc, glucose; glyc, glyceryl; MeOH, methanol; Pyr, pyruvate; Un 1, unknown 1 (7.19-7.38, br); Un 3, unknown 3 (7.71, m). Values in brackets correspond to compound chemical shifts.

Table 6.2. List of metabolites/resonances in whole blood plasma and plasma lipid extracts selected by variables selection as important for discrimination of CD and T21 from controls, in both 1st and 2nd trimesters. [a] Chemical shifts shown correspond to signals used for integration; s: singlet, d: doublet, t: triplet, q: quartet, m: multiplet, br: broad. [b] E.S.: effect size determined as described in reference (Berben et al., 2012). [c] Significance level 95% (p-value < 0.05). Only values with statistical relevance or lower uncertainties are specified, otherwise only qualitative information is provided in the form of arrows. *: observations noted only in the 2nd T; †: observations noted only the T21 subgroup. n.d.: not determined.

Compound	Chemical shift/ppm[a]	Variation (E.S.[b], p-value[c])			
		1st trimester		2nd trimester	
		CD *vs.* controls	T21 *vs.* controls	CD *vs.* controls	T21 *vs.* controls
Whole blood plasma					
CH_3 lipids HDL	0.78 -0.86 (br)	↓ -0.86±0.67, 1.3×10^{-2}	↓ -1.10±0.84, 8.8×10^{-3}	↓	↓
$(CH_2)_n$ lipids HDL	1.20-1.25 (br)	↓ -0.68±0.66, 4.9×10^{-2}	↓	↓	↓ -0.56±0.60, 2.2×10^{-2}
CH_3 lipids LDL+VLDL	0.86-0.90 (br)	-	-	↑*	↑
$(CH_2)_n$ lipids LDL+VLDL	1.25-1.30 (br)	-	-	↑*	↑
CH_3 HDL/LDL+VLDL		↓ -0.73±0.67, 4.7×10^{-2}	↓	↓	↓
Alanine	1.46 (d)	-	-	↑*	↑
$C\underline{H_2}CH_2CO$ lipids	1.52-1.60 (br)	↑	↑	-*	↑†
Acetate	1.91 (s)	↑	↑	↓ -0.51±0.47, 1.2×10^{-2}*	↓
Acetone	2.22 (s)	↑	↑	-	-
β-hydroxybutyrate	2.39 (m)	↑ 0.69±0.66, 2.6×10^{-3}	↑ 0.70±0.82, 1.8×10^{-2}	↑	-†
Pyruvate	2.36 (s)	↑ 0.81±0.67, 3.0×10^{-2}	↑ 1.22±0.85, 4.6×10^{-2}	↑	↑
Citrate	2.51 (d)	↑	↑	↑ 0.51±0.47, 3.3×10^{-2}	↑
Albumin-lysyl	2.92-3.03 (br)	-	-	-	↑†
Creatine	3.03 (s)	↑	↑	-*	-
$N(CH_3)_3$ choline HDL	3.20-3.23 (br)	↓ -1.06±0.69, 3.9×10^{-3}	↓ -1.28±0.86, 5.1×10^{-3}	↓	↓
Proline	3.34 (br)	↓ -0.78±0.67, 1.7×10^{-2}	↓ -1.02±0.84, 1.8×10^{-2}	↓	-†
Methanol	3.36 (s)	↓ -0.98±0.68, 3.0×10^{-3}	↓ -1.01±0.84, 2.3×10^{-2}	-*	↑†
Creatinine	4.05 (s)	-	-	↑ 0.60±0.47, 1.4×10^{-2}*	-†
Glucose	5.23 (d)	↑	↑	↑	↑ 0.47±0.60, 3.0×10^{-2}
Urea	5.77 (br)	-	-	↓ -0.81±0.48, 6.7×10^{-4}*	↓ -0.87±0.61, 2.6×10^{-3}
Unknown 1	7.19-7.38 (br)	-	-	-	↑†

Table 6.2. (cont.)

Compound	Chemical shift/ppm[a]	Variation (E.S.[b], p-value[c]) 1st trimester		2nd trimester	
		Controls *vs.* CD	Controls *vs.* T21	Controls *vs.* CD	Controls *vs.* T21
Plasma lipid extracts					
Fatty acids					
CH_3	0.89 (t)	n.d.	↓ -0.99±0.67, 1.3×10^{-2}	n.d.	↓
$(CH_2)_n$	1.25 (br)	n.d.	↓	n.d.	↓ -0.56±0.54
$=CHCH_2(C\underline{H}_2)_n$	1.26-1.40 (br)	n.d.	↓	n.d.	-*
$C\underline{H}_2CH=$ 18:2	2.05 (q)	n.d.	↓ -0.86±0.66, 1.7×10^{-2}	n.d.	-*
$C\underline{H}_2CH=$ 20:4	2.10 (q)	n.d.	↓	n.d.	↓
CH_2CO	2.28-2.36 (m)	n.d.	↓	n.d.	↓
$C\underline{H}_2CO$ 22:6	2.37 (t)	n.d.	↑	n.d.	↓ -0.94±0.56, 8.5×10^{-4}*
$=CHC\underline{H}_2CH=$ 18:2	2.77 (t)	n.d.	↓	n.d.	↑*
$=CHC\underline{H}_2CH=$ 20:4	2.79-2.86 (m)	n.d.	↓	n.d.	↓ -0.83±0.55, 3.5×10^{-3}
HC=CH 18:2	5.29-5.35 (m)	n.d.	↓	n.d.	-*
HC=CH 20:4	5.35-5.43 (m)	n.d.	↓	n.d.	↓ -0.79±0.55, 6.5×10^{-3}
Phospholipids					
$N(CH_3)_3$ choline SM	3.31 (br)	n.d.	↓	n.d.	-*
$N(CH_3)_3$ choline PC	3.34 (br)	n.d.	↓ -1.36±0.70, 2.1×10^{-4}	n.d.	↓ -1.11±0.57, 5.5×10^{-5}
$C\underline{H}_2N(CH_3)_3$ choline total	3.78 (br)	n.d.	↓ -1.59±0.72, 8.7×10^{-5}	n.d.	↓ -1.30±0.58, 2.5×10^{-5}
Glyceryl-$C3H_2$	3.95 (m)	n.d.	↓ -1.53±0.71, 1.1×10^{-4}	n.d.	↓ -1.14±0.57, 1.7×10^{-4}
$POCH_2$	4.34 (br)	n.d.	↓ -0.83±0.66, 6.8×10^{-3}	n.d.	↓ -1.36±0.59, 3.0×10^{-5}
Glyceryl-$C1H_2$	4.40 (br)	n.d.	↓ -1.08±0.67, 4.2×10^{-3}	n.d.	↓ -0.79±0.55, 6.1×10^{-3}
Unknown 3	7.71 (m)	n.d.	-	n.d.	↑ 1.08±0.57, 3.9×10^{-4}*

In an attempt to improve samples classification, multivariate analysis was performed using a concatenated matrix of 2nd T maternal blood plasma and urine collected for common subjects (22 controls, 26 CDs, 12 T21). This led, however, to models with comparable robustness as those considering plasma or urine (Diaz et al., 2013b) alone (Table 6.1), identifying metabolite variations (not shown) in general agreement with those provided by single-biofluid models. However, in the smaller concatenated matrix (particularly diminished for controls), the urea and creatinine variations (shown in plasma-only) were absent, thus demonstrating the importance of using larger cohort sizes. Metabolite variations were then used to identify intra- and inter-biofluid metabolic correlations describing CD cases in the 2nd T. The resulting correlations network (Figure 6.5) identified one correlation cluster within plasma (A) and two main clusters in urine (B and C). Cluster A seems to reflect the lipid metabolism disturbances, unveiling a negative correlation between plasma creatine and lipids and a positive correlation between ketone body acetoacetate and circulating lipids. Cluster B links the three excreted ketone bodies, all correlated to acetate; ketone bodies also correlate inversely to excreted formate (node 40) and directly to plasma acetoacetate (node 9). Cluster C seems to reflect deviations in Krebs cycle intermediates (2-ketoglutarate, fumarate), interlinked to glucose, pyruvate and taurine. Apparent additional correlations are observed in urine, namely between 1) excreted glucose, sucrose and allantoin (nodes 30, 33, 31); 2) excreted 4-hydroxyhippurate, indoxyl sulphate and phenylacetylglutamine (nodes 35-37), with the former correlating inversely to creatinine (node 29) and this, in turn, to trigonelline (node 41). Notably, the relatively large variation in excreted 3-methyl-histidine (node 39) does not show relevant correlations.

6.1.3. Proposed metabolic interpretation of plasma changes

In the 1st T, plasma of CD and T21 groups share decreased levels of HDL lipoproteins and increased levels of ketone bodies (β-hydroxybutyrate and acetone), suggesting increased β-oxidation. Increases in other metabolites connected to energy metabolism (namely glucose, pyruvate and citrate) seem to reflect their underuse in glycolysis, pyruvate oxidation and Krebs cycle, respectively, indicating an energy shift from glucose to lipids catabolism. In T21 cases, 18:2 FA and PC levels are particularly decreased, possibly reflecting their role as HDL constituents. In addition, proline, methanol and creatine add to an early CD-specific signature, although no clear biochemical interpretation may be advanced at this stage.

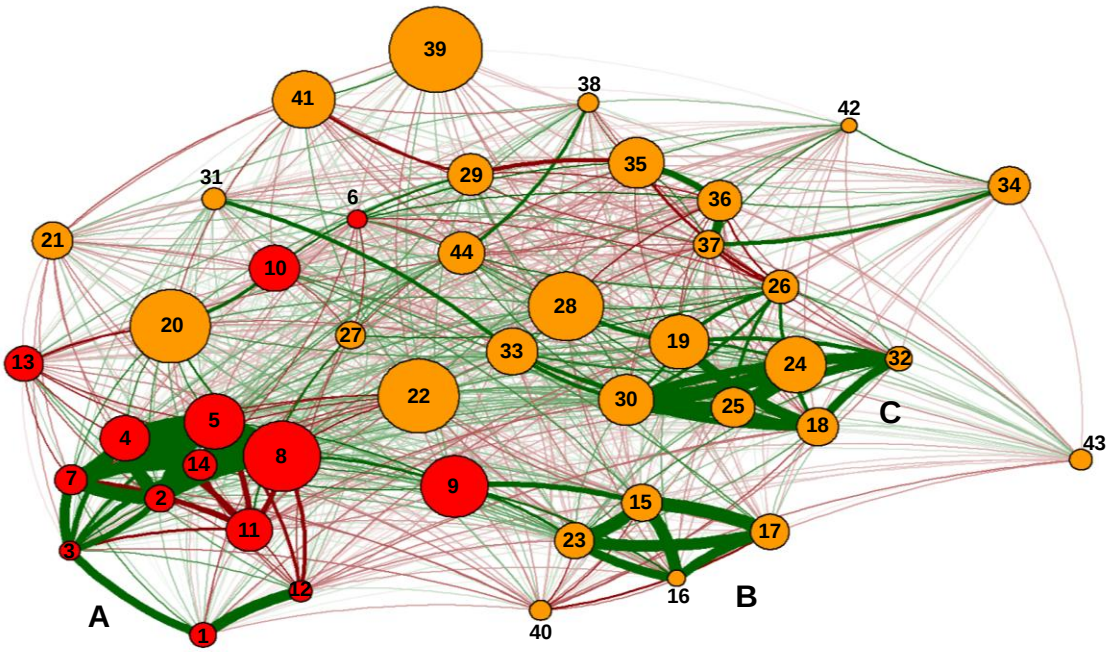


Figure 6.5. Correlation network of maternal plasma and urine metabolites observed to change in 2nd trimester CD cases compared to controls, based on Pearson's correlation coefficients. Node colours indicate biofluid type (red for blood plasma and orange for urine) and node size reflects effect size of variation in CDs, compared to controls. Green and red lines correspond to positive and negative correlations, respectively, and line thickness reflects correlation coefficients magnitude. 1- CH_3 lipids HDL; 2- CH_3 lipids LDL+VLDL; 3- $(CH_2)_n$ lipids HDL; 4- $(CH_2)_n$ lipids LDL+VLDL; 5- $C\underline{H_2}CH_2CO$ lipids; 6- acetate; 7- $CH_2C{=}C$ lipids; 8- CH_2CO lipids; 9- acetoacetate; 10- citrate; 11- creatine; 12- $N(CH_3)_3$ choline HDL; 13- lactate; 14- HC=CH lipids LDL+VLDL; 15- acetone; 16- acetate; 17- β-hydroxybutyrate; 18- pyruvate; 19- 2-ketoglutarate; 20- citrate; 21- dimethylamine; 22- *scyllo*-inositol; 23- acetoacetate; 24- taurine; 25- unknown 5 (3.48-3.51 ppm, m); 26- guanidoacetate; 27- creatine; 28- unknown 6 (3.96 ppm, s); 29- creatinine; 30- glucose; 31- allantoin; 32- fumarate; 33- sucrose; 34- 4-hydroxylphenylacetate; 35- 4-hydroxyhippurate; 36- indoxyl sulphate; 37- phenylacetylglutamine; 38- histidine; 39- 3-methyl-histidine; 40- formate; 41- trigonelline; 42- malonate; 43- tartarate; 44- choline.

In the 2nd T, the observation of different metabolic profiles for CD and T21 groups, compared to 1st T, suggests that a dynamic metabolic evolution occurs in CD cases, rather than a simple enhancement of the metabolic impact registered early on. The metabolic pattern of enhanced β-oxidation is maintained (with the exception of acetone variation), however, 2nd T plasma includes novel changes, namely increased LDL+VLDL (possibly as compensation for HDL depletion) and decreased urea. The latter, together with an increasing (although not statistically correlated) tendency in creatinine, may indicate changes in fetal renal function, a condition often observed in CD-affected fetuses (Creasy et al., 2009) and here seen to acquire particular emphasis in the 2nd T. The lower urea levels may result from up-regulation of creatine phosphorylation and/or down regulation of creatine hydrolysis into urea and sarcosine; the latter was supported by STOCSY results (not shown) directly correlating urea and

sarcosine. Second trimester T21 cases show a similar behavior to global CD, with additional increasing tendencies for methanol and albumin; the latter may relate to the lower HDL levels, since an inverse albumin/HDL relationship was observed in healthy pregnancies (Pinto et al., 2014b). Additional apparently T21-specific deviations comprised decreases in ω-3 FA 22:6 and ω-6 FA 20:4, known for their mediation in anti-inflammatory and pro-inflammatory processes (Wall et al., 2010), respectively. The somewhat larger depletion in the former may indicate a more significant response to inflammation, accompanying 2nd T T21.

6.1.4. Conclusions

These results partially confirm previously reported changes in deproteinized 1st T maternal blood sera of T21 and T18 cases (Bahado-Singh et al., 2013b, 2013c) and add information regarding additional metabolites (proline, methanol) and lipid metabolism in CD and T21. Other alterations previously observed in T21 cases (Bahado-Singh et al., 2013b) were not noted here probably due to metabolite low abundance and resonance overlap with macromolecules in whole plasma. In addition, a comparison between 1st and 2nd T was performed in this work, enabling a time course characterization of CD (and T21) to be carried out and in tandem analysis of plasma and lipid extracts unveiled specific changes in lipid metabolism. Indeed, changed phospholipid levels had been reported for T21 (Pinto et al., 2014d) and, here, a different involvement of HDL and LDL+VLDL lipoproteins in β-oxidation was identified, in addition to a fine interplay between 18:2, 22:6, 20:4 FA. Plasma/urine correlations studies unveiled possible CD-specific intra- and inter-biofluid metabolic relationships, confirming main changes noted here for plasma and previously for urine (Diaz et al., 2013b) and unveiling possible relationships relating to plasma creatine and lipids (particularly LDL+VLDL); excreted ketone bodies and acetate/formate; allantoin and sugar metabolism; and gut-microflora metabolites (e.g., 4-hydroxyhippurate, phenylacetylglutamine, 4-hydroxyphenylace-tate).

In future work, the putative metabolic signatures advanced for 1st T and 2nd T maternal plasma in CD and T21 require validation in larger cohorts and demonstration of biochemical explanations, before clinical applications may be envisaged. Additionally, the issues of metabolic signature specificity in relation to T21, T18 and T13 requires further investigation.

6.2. Analysis of plasma phospholipids by Hydrophilic Interaction Liquid Chromatography-Mass Spectrometry (HILIC-LC/MS)

The results presented below were published as part of a short communication with reference: Pinto, J.; Maciel, E.; Melo, T. S.; Domingues, M. R. M.; Galhano, E.; Pita, C.; Almeida, M. do C.; Carreira, I. M.; Gil, A. M. Maternal Plasma Phospholipids Are Altered in Trisomy 21 Cases and prior to Preeclampsia and Preterm Outcomes. *Rapid Communications in Mass Spectrometry* **2014**, *28*, 1635–1638. In this subchapter only the results obtained for 2nd T T21 will be described. The brief state of the art was described in subchapter 6.1. Sample characteristics can be found in Chapter 2 (Table 2.6).

Abstract

Rationale: To search for changes in phospholipids in 2nd T maternal plasma, in connection to fetal T21, aiming at finding novel markers. Methods: Plasma was collected from pregnant women at 16-19 g.w. defining two groups: controls (n=14) and T21 (n=5). The PL extracts of the pooled samples were analysed by hydrophilic interaction liquid chromatography mass spectrometry (HILIC-LC/MS) and subsequent multivariate and univariate analysis. Results: Results showed that T21 has a significant impact on 2nd T maternal plasma PL: changes in the levels of LPC, ether-linked PC and PE are consistent with previously suggested oxidative processes related to T21. Apparent specific features included decreases in two SM and several PC. Conclusions: Results indicated that maternal PL metabolism is exquisitely sensitive to fetal T21 and the apparent specificity of PL profiles enable future uses for complementary diagnosis to be envisaged.

6.2.1. Phospholipid profiling of trisomy 21 in 2nd trimester

Typical positive-ion HILIC-LC/MS records (Figure 6.6) of pooled 2nd T control plasma showed peaks characteristic of different PL classes at average RT 9.5–10.8 min (PE), 29.5–32.4 min (PC), 34.0–35.9 min (SM) and 38.5–39.8 min (LPC).

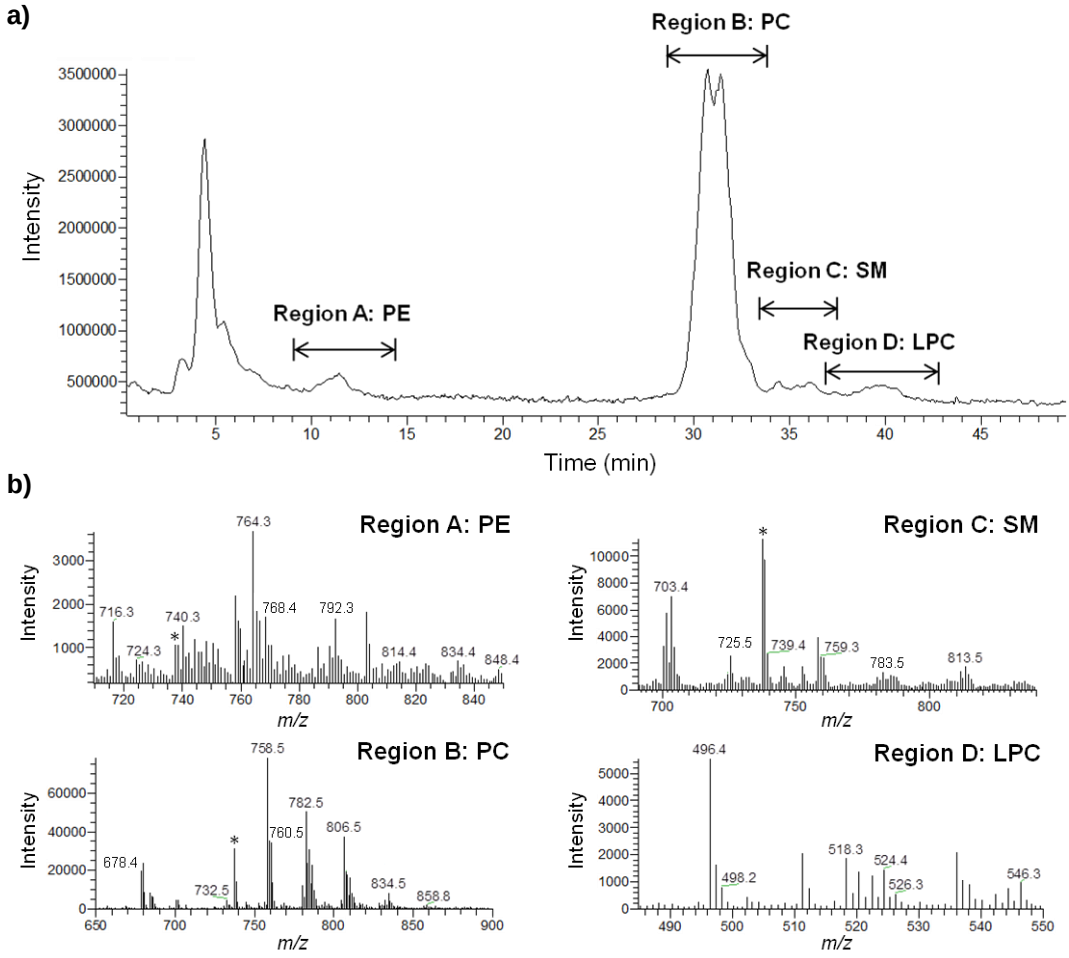


Figure 6.6. Positive-ion HILIC-LC/MS analysis of a pooled plasma sample from control pregnant women. a) Total ion chromatogram (TIC) and b) LC-MS spectra for each PL class: region A: phosphatidylethanolamines (PE); region B: phosphatidylcholines (PC); region C: sphingomyelins (SM); and region D: lyso PC (LPC).

When visually comparing the corresponding MS records obtained for control and disease pools (see Figure 6.7a, for PC species), some differences could be observed for T21 (arrows in Figure 6.7a). A more objective analysis of such complex datasets requires the use of multivariate analysis, used here to probe for changes to be further confirmed rather than for classification (due to the low sample numbers). Preliminary PLS-DA (Figure 6.7b, left) showed separation of T21 from controls with a good predictive power (Q^2= 0.574). Loadings inspection (Figure 6.7b, right) enabled a total of 34 features to be identified with VIP $\geq$ 1 and $p < 0.05$ for T21. The most relevant features (with higher VIP and lower p value,

Table 6.3) included significant decreases in SM (d18:1/18:1) (Figure 6.7c) and unknown species with *m/z* 842.3 (unquantifiable in T21 samples).

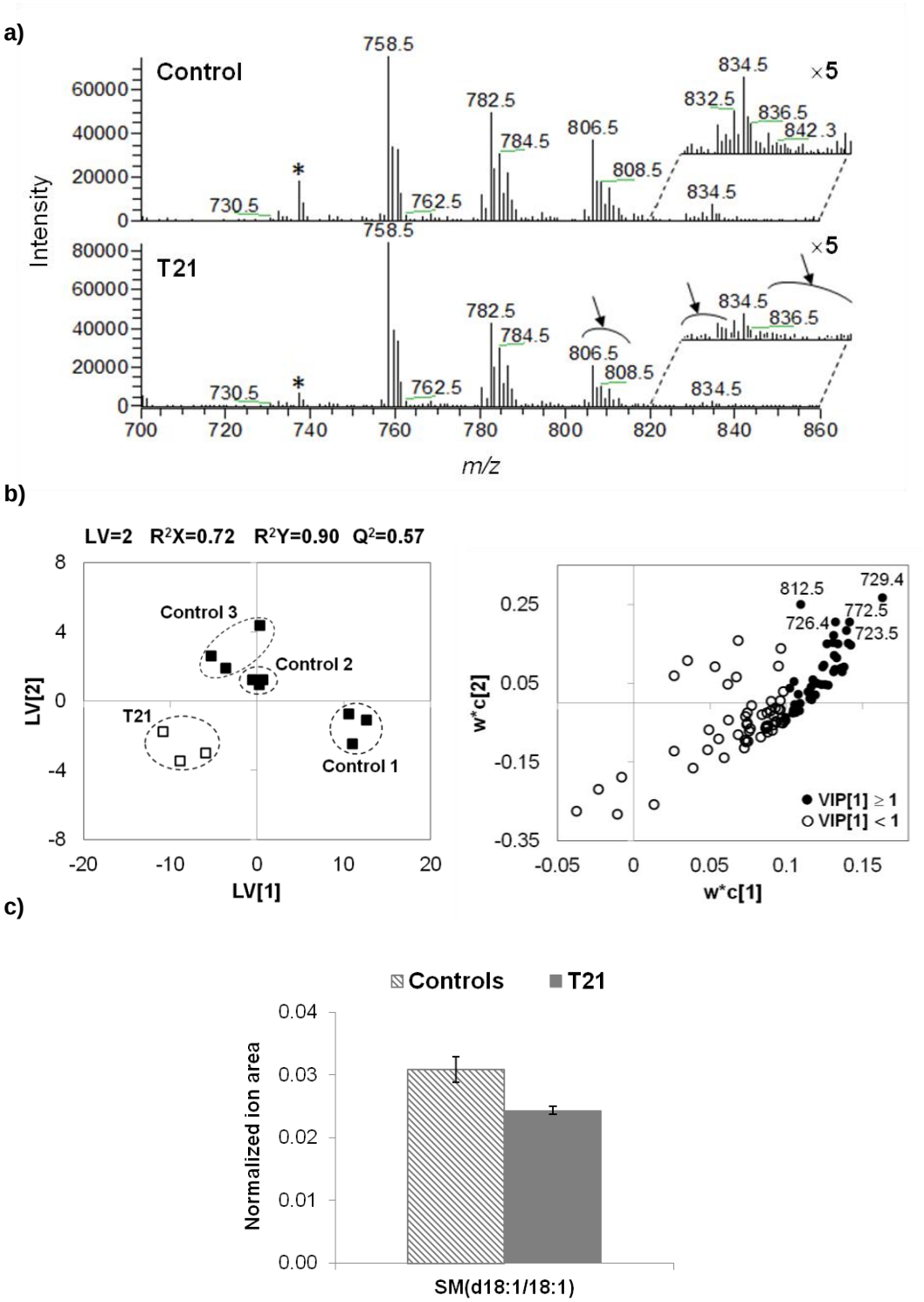


Figure 6.7. a) MS record of PC TIC region (average RT 29.5 – 32.4 min) of pooled plasma samples of pregnant women: healthy throughout pregnancy (control) and carrying fetuses affected by trisomy 21; b) PLS-DA scores scatter plot and loadings plot of controls *vs.* T21; c) Bar plot of normalized ion area of a selected PL for control vs. T21 samples pools. * solvent impurity.

Table 6.3. List of PL species varying most significantly between T21 *vs.* controls samples. Abbreviations: *m/z*: mass-to-charge ratio; RT: retention time; VIP: variable importance to projection. *unknown compound, † signal observed in more than one disorder (unspecific).

m/z	RT (min)	Identification	VIP[1]	p-value (<0.05)	Relative variation (%)	Effect size
729.4	35.0	SM(d18:1/18:1)	1.64	1.87×10^{-5}	-21.2 ± 2.6	-3.23 ± 0.90
842.3	30.9	Unknown * †	1.44	-		-
542.3	39.3	LPC(20:5) †	1.42	2.69×10^{-4}	-46.8 ± 8.4	-2.23 ± 0.76
762.5	31.3	PC(34:0)	1.39	1.20×10^{-3}	-33.7 ± 7.0	-2.15 ± 0.75
522.4	39.1	LPC(18:1) †	1.38	4.68×10^{-4}	-41.3 ± 7.7	-2.10 ± 0.75
675.4	35.9	SM(d18:1/14:0)	1.37	5.38×10^{-4}	-34.4 ± 6.9	-2.05 ± 0.74
744.5	31.6	PC(O-34:2)	1.33	6.37×10^{-4}	-26.3 ± 5.3	-2.03 ± 0.74
856.5	29.8	PC(O-42:2) (and PC(42:9))	1.32	6.97×10^{-4}	-33.0 ± 6.9	-1.92 ± 0.73
752.3	9.7	PE(O-38:5)	1.32	6.24×10^{-3}	-24.2 ± 5.8	-1.90 ± 0.72
832.5	29.9	PC(O-40:0) (and PC(40:7))	1.32	3.64×10^{-2}	-45.3 ± 23.4	-1.91 ± 0.72
544.4	38.5	LPC(20:4) †	1.29	9.84×10^{-4}	-42.6 ± 9.2	-2.06 ± 0.74
858.5	30.1	PC(O-42:1) (and PC(42:8))	1.28	1.62×10^{-3}	-60.7 ± 15.2	-1.80 ± 0.71
808.5	30.1	PC(38:5)	1.27	7.92×10^{-3}	-23.6 ± 6.1	-1.79 ± 0.71
808.5	10.7	PE(O- 42:5)	1.25	3.44×10^{-3}	-78.1 ± 22.3	-1.76 ± 0.71
744.4	10.3	PE(36:2)	1.24	1.58×10^{-2}	-35.0 ± 10.1	-1.72 ± 0.71
836.5	29.9	PC(40:5)	1.21	1.47×10^{-2}	-33.8 ± 9.8	-1.71 ± 0.70
520.4	39.4	LPC(18:2) †	1.20	2.05×10^{-3}	-28.5 ± 6.7	-1.60 ± 0.69

6.2.2. Proposed metabolic interpretation of phospholipid changes

Our results show, for the first time, that T21 has a significant impact on 2nd T maternal plasma PL. The changed LPC, along with ether-linked PC and PE species, are known to relate to lipid oxidation processes, seemingly characterizing fetal T21, as previously suggested with basis on proteomics (Perluigi et al., 2011). Furthermore, decreases in two SM are apparently specific of T21 (compared with pre-PTD and pre-preeclampsia in Chapter 8, section 8.2), along with a number of decreasing PC. LPC species are usually formed due to enzymatic digestion by phospholipase A2, thus leading to LPC bearing predominantly saturated fatty acyl chains. Interestingly, in this work, although LPC with saturated fatty acyl chains were detected within the plasma lipidome (Figure 6.8), the LPC species that showed significant variations between controls and T21 were those bearing unsaturated PC. LPC comprising unsaturated fatty acyl chains have been identified before in plasma (Quehenberger et al., 2010) and may probably arise from the action of phospholipase A1 or to oxidative cleavage by metal-catalyzed radical oxidation.

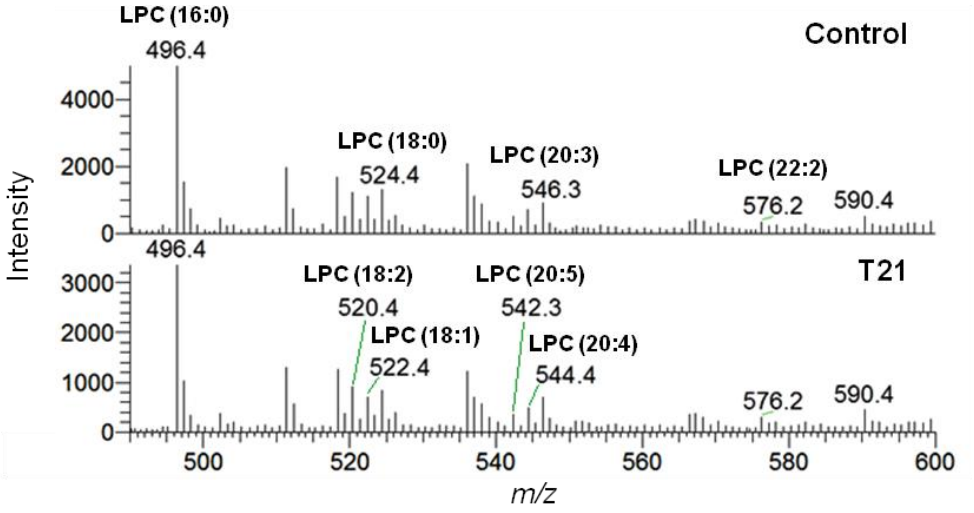


Figure 6.8. MS record of LPC TIC region (average RT 38.5–39.8 min) of pooled plasma samples of pregnant women: healthy throughout pregnancy (control) and carrying fetuses affected by T21.

These results indicated that maternal PL metabolism is exquisitely sensitive to fetal T21, although consideration of larger cohorts is necessary for adequate statistical validation to be carried out. The PL marker profiles observed seem specific enough to enable future uses as complementary diagnosis of T21 to be envisaged, for improved prenatal health management.

Chapter 7. Second trimester maternal plasma for prediction of gestational diabetes mellitus

The results presented in this chapter were fully submitted for publication in the reference below and are here adapted slightly onto the form of a thesis chapter, for the sake of clarity:

Pinto, J.; Almeida, L. M.; Martins, A. S.; Duarte, D.; Barros, A. S.; Galhano, E.; Pita, C.; Almeida, M. C.; Carreira, I. M.; Gil, A. M. Prediction of gestational diabetes through NMR metabolomics of maternal blood. *Journal of Proteome Research* **2015**, DOI: 10.1021/acs.jproteome.5b00260.

Abstract

Metabolic biomarkers of pre- and post-diagnosis GDM were sought, using NMR metabolomics of maternal plasma and corresponding lipid extracts. Metabolite differences between controls and disease were identified through multivariate analysis of variable selected ^{1}H NMR spectra. For post-diagnosis GDM, Partial Least Squares (PLS) regression identified metabolites with higher dependence on normal gestational age evolution. Variable selection of NMR spectra produced good classification models for both pre- and post-diagnostic GDM. Pre-diagnosis GDM was accompanied by cholesterol increase and minor increases in lipoproteins (plasma), fatty acids and triglycerides (extracts). Small metabolite changes comprised variations in glucose (up regulated), amino acids, betaine, urea, creatine and metabolites related to gut microflora. Most changes were enhanced upon GDM diagnosis, in addition to newly observed changes in some small compounds. GDM prediction seems possible exploiting multivariate profile changes rather than a set of univariate changes. Post-diagnosis GDM is successfully classified using a 26-resonance plasma biomarker. Plasma and extracts display comparable classification performance, the former enabling direct and more rapid analysis. Results and putative biochemical hypotheses require further confirmation in larger cohorts of distinct ethnicities.

Brief state of the art

Given the important risks involved with GDM pregnancies and infants, there is still scope for devising earlier and more complete biomarkers of the condition onset and development. In recent years, metabolomic studies have searched for metabolic biomarkers of GDM by profiling maternal biofluids (blood serum/plasma and urine) and fetal/infant samples (amniotic fluid, umbilical cord blood, newborn meconium and urine), to obtain fuller descriptions of GDM impact on maternal and fetal metabolisms, as reviewed recently (Huynh et al., 2014). Some of these studies have sought biomarkers detectable prior to clinical diagnosis, so that women at risk of developing GDM may be identified in time to allow improved disease management. A large number of studies may be found in relation to post-diagnosis cases, an initial GC-MS study revealing an increase in maternal serum total FA and several specific FA (e.g., linoleic, arachidonic, docohexaenoic) from controls to slight hyperglycaemia and to 3rd trimester GDM, with differences in palmitoleic and docosahexenoic acid levels found between hyperglycaemia and GDM (Chen et al., 2010). A subsequent study showed higher serum levels of TG, 3-hydroxybutyrate, and Ala, Pro, and branched-chain amino

acids (BCAA), and lower 1,5-anhydroglucitol levels, in women with higher fasting glucose levels (Scholtens et al., 2014). GDM impact on lipid metabolism was confirmed by LC- and GC-MS of maternal plasma, where several lysophospholipids, taurine-bile acids and long-chain polyunsaturated FA derivatives were found discriminative of GDM and indicative of low-grade inflammation and altered redox-balance (Dudzik et al., 2014). To our knowledge, no Nuclear Magnetic Resonance (NMR) metabolomics characterisation (a more holistic approach, compared to MS) of maternal blood related to post-diagnosis GDM has been reported.

Regarding the search for metabolic alterations in maternal plasma preceding GDM diagnosis, NMR of a small cohort detected decreases in TMAO and betaine, interpreted as changes in choline-related homocysteine-methionine conversion and possible alterations in renal function (Diaz and Pinto et al., 2011). Recently, GC-MS of maternal sera unveiled early increases in itaconic acid and *cis*-aconitate, possibly due to inflammation (de Seymour et al., 2014), and a subsequent LC-MS study unveiled increases in anthranilic acid, Ala, Glu, Ser and alantoin, and decreased creatinine (Bentley-Lewis et al., 2015). The profiling of maternal urine, amniotic fluid, cord serum, and newborn urine and meconium has also been performed, a detailed description of these studies being found in Chapter 1 (section 1.4.1).

The present paper presents a NMR metabolomics study of maternal plasma and corresponding lipid extracts obtained for a group of pregnant women without clinical signs of the disease but who 2-21 g.w. later developed GDM (n=32, pre-diagnosis group) and a group of pregnant women at the time of GDM diagnosis (n=12, post-diagnosis group), in comparison to controls. This adds to a previous smaller cohort NMR study of pre-diagnosis GDM (n=14) reported by our own group (Diaz and Pinto et al., 2011) and provides a first NMR study of post-diagnosis GDM, to the best of our knowledge. In the latter case, GDM and control cases differed in average gestational age and, hence, the metabolites related to the normal underlying pregnancy evolution were identified prior to determining the post-diagnosis GDM metabolic profile. Furthermore, the complementary information provided by whole plasma and lipid extract analysis is discussed and sample sensitivity for GDM compared.

7.1. Analysis of whole plasma and lipid extracts

Figure 7.1 shows the average ^{1}H CPMG NMR spectra of blood plasma for controls, pre- and post-diagnosis GDM groups. Some changes are visible for the post-diagnosis GDM group, compared to controls (arrows in Figure 7.1c), namely in lipids $(CH_2)_n$, Ala,

$N(CH_3)_3$ choline of PL and unknown resonance at δ 7.15-7.35, whereas no obvious changes are noted for the pre-diagnosis GDM group (Figure 7.1b).

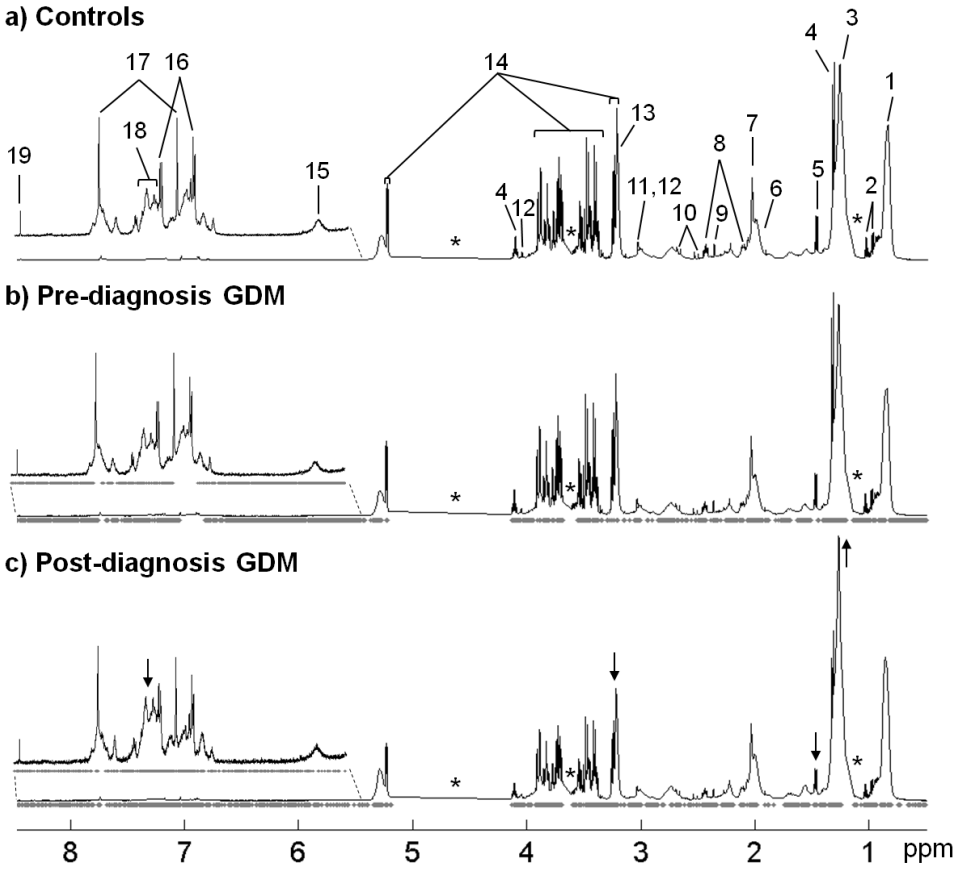


Figure 7.1. Average 500 MHz CPMG ^{1}H NMR spectra of plasma of a) controls, b) pre- and c) post-diagnosis GDM groups, with indication of spectral variables (data points) selected in the corresponding PLS-DA models (grey dots under spectra). 1- CH_3 lipids, 2- Val, 3- $(CH_2)_n$ lipids, 4- lactate, 5- Ala, 6- $CH_2C{=}C$ lipids, 7- N-acetyl glycoproteins, 8- Gln, 9- pyruvate, 10- citrate, 11- creatine, 12- creatinine, 13- $N(CH_3)_3$ choline of PL, 14- glucose, 15- urea, 16- Tyr, 17- His, 18- unknown (δ 7.15-7.35), 19- formate. Arrows indicate visible alterations. *: cut off spectral regions corresponding to water and ethanol peaks.

Regarding lipid extracts (standard spectra, Figure 7.2), small increases are seen in several lipid resonances, from controls to pre-diagnosis and to post-diagnosis GDM (Figure 7.2b), the spectra giving a detailed account of the different components of plasma lipids and lipoproteins, namely FA (in particular, 18:2 and 20:4), cholesterol (free and esterified), TG and PL (in particular, PC and SM).

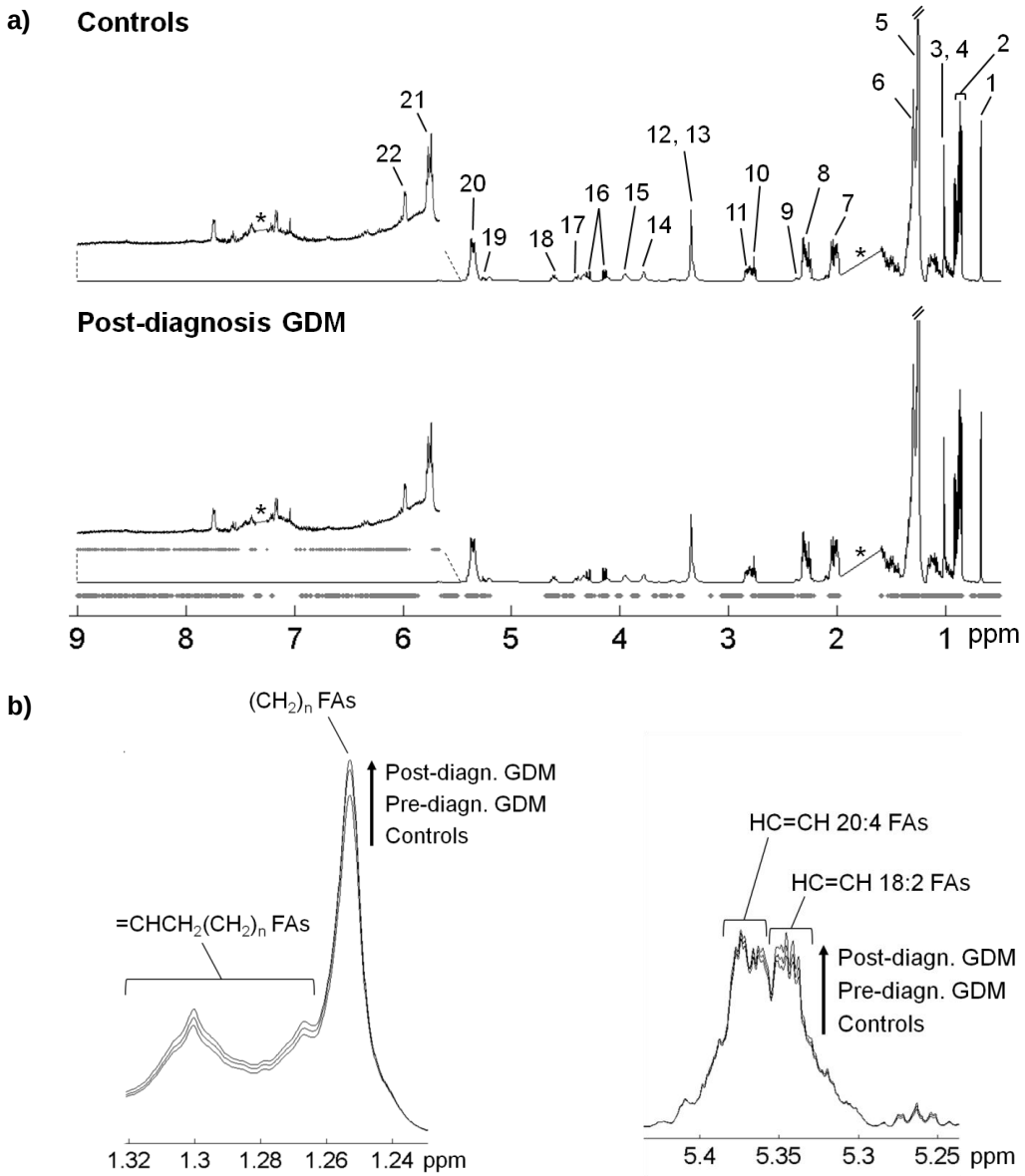


Figure 7.2. a) Average (n=15) 500 MHz ^{1}H NMR spectra of plasma lipid extracts of controls (top) and post-diagnosis GDM (bottom), with indication of spectral variables selected in the corresponding PLS-DA model (grey dots under bottom spectrum); b) spectral expansions of the δ 1.23-1.32 (left) and δ 5.24-5.44 (right) regions of overlaid average spectra of controls, pre- and post-diagnosis GDM. Assignment: 1- $C18H_3$ FC and EC, 2- $C26H_3$ and $C27H_3$ FC and EC, 3- $C19H_3$ FC, 4- $C19H_3$ EC, 5- $(CH_2)_n$ FA, 6- $=CHCH_2(C\underline{H}_2)_n$ FA, 7- $CH_2CH=$ FA, 8-CH_2CO FA, 9-CH_2CO 22:6 FA, 10- $=CHC\underline{H}_2CH=$ 18:2 FA, 11- $=CHC\underline{H}_2CH=$ 20:4 FA, 12- $N(CH_3)_3$ choline SM, 13- $N(CH_3)_3$ choline PC, 14- $CH_2N(CH_3)_3$ PC and SM, 15- glyceryl-$C3H_2$ of PL, 16- glyceryl-C1,3H and –C1,3H' of TG, 17- glyceryl-$C1H_2$ and CH_2-PO PL, 18- C3H EC, 19- glyceryl-C2H TG, 20- HC=CH FA, 21- =CH-CHOH SM, 22- Unknown 3 (δ 5.91, br).

Pre-diagnosis GDM vs. controls

Multivariate analysis was performed for comparison of controls and pre-diagnosis GDM groups, matched for gestational age (Table 2.5, Section 2.1.2, page 67). PLS-DA of plasma spectra produced non-robust models for both CPMG and diffusion-edited spectra (not shown), as revealed by the corresponding low predictive power (Q^2), %

spec. and % sens. given by MCCV (Table 7.1). However, upon variable selection, a good model was obtained for CPMG spectra (Figure 7.3a, left), whereas a weaker performance was noted when considering the diffusion-edited spectra (Figure 7.3b, left and Table 7.1). The lower sensitivities (lower true positive rates, TPR) express the fact that several pre-diagnosis GDM cases are, in these models, wrongly classified as controls (Figure 7.3a,b left). The grey dots under the CPMG spectrum in Figure 7.1b represent the data points selected as main contributors for pre-diagnosis GDM group separation from controls in Figure 7.3a. Analysis of the corresponding loadings (Figure 7.4a) and signal integration enabled peak (and metabolite) variations to be quantified and their univariate statistical relevance assessed (Table 7.2).

Table 7.1. MCCV parameters obtained when considering original (full dataset) and variable selected ^{1}H NMR spectra. LV: no. of latent variables; Q^2: most frequent value of predictive power obtained by MCCV; CR: classification rate; sens: sensitivity; spec: specificity; %VS: percentage of variables selected from the original spectra. Values in bold refer to robust PLS-DA models.

Subject group	Original spectra					Variable selection					
	LV	Q^2	CR	Sens	Spec	LV	Q^2	CR	Sens	Spec	% VS
Controls vs. pre-diagn. GDM											
Whole plasma /CPMG	1	0.1	68%	53%	77%	**2**	**0.5**	**81%**	**69%**	**89%**	**36%**
Whole plasma /Diffusion-ed.	1	0.3	73%	60%	81%	1	0.2	72%	57%	82%	38%
Lipid extracts	1	0.3	64%	63%	64%	**1**	**0.5**	**83%**	**75%**	**90%**	**33%**
Controls vs. post-diagn. GDM											
Whole plasma /CPMG	1	0.4	86%	43%	96%	**1**	**0.6**	**90%**	**59%**	**97%**	**32%**
Whole plasma /Diffusion-ed.	**1**	**0.5**	**93%**	**80%**	**97%**	**1**	**0.7**	**92%**	**77%**	**96%**	**38%**
Lipid extracts	**4**	**0.6**	**84%**	**80%**	**87%**	**2**	**0.7**	**79%**	**69%**	**87%**	**29%**

The results show that, several weeks prior to GDM diagnosis (Table 2.5, Section 2.1.2, page 67), plasma shows a changed profile in low M_w metabolites, comprising changes with $p < 0.05$ for five low M_w compounds (Val↑, Pyr↑, Pro↓, urea↓ and 1,5-anhydroglucitol↓), in tandem with variation tendencies in Gln (↓), creatine (↓), dimethyl sulfone (↓), trimethylamine N-oxide (TMAO) (↓), betaine (↑), lactate (↑) and glucose (↑) (Table 7.2).

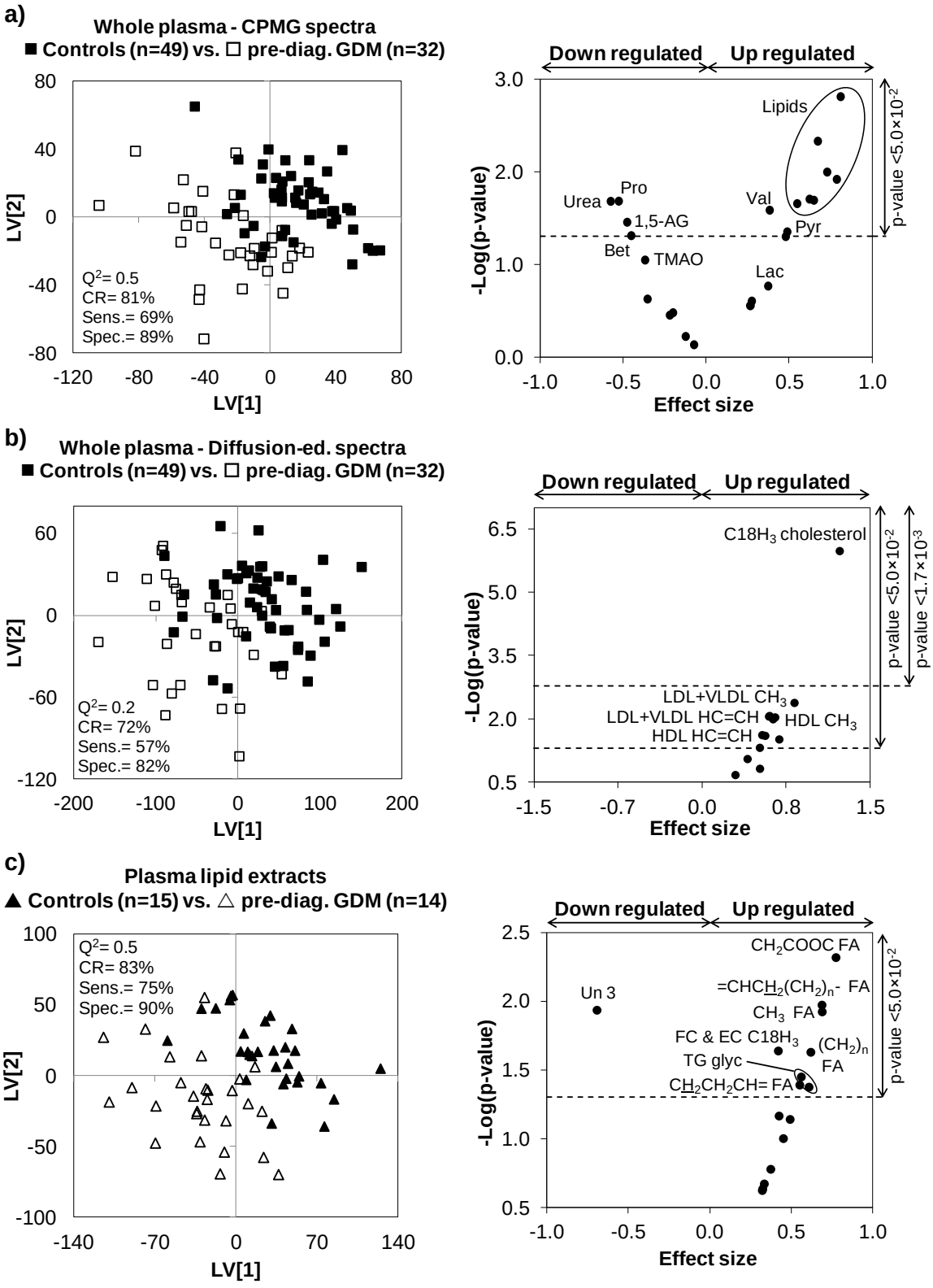


Figure 7.3. PLS-DA scores scatter plots and Volcano plots (effect size *vs.* – Log (p-value)) of a) plasma CPMG ^{1}H NMR spectra of controls (■, n=49) *vs.* pre-diagnosis GDM (□, n=32) (R^2X 0.29, R^2Y 0.58), b) plasma diffusion-edited ^{1}H NMR spectra of controls (■, n=49) *vs.* pre-diagnosis GDM (□, n=32) (R^2X 0.49, R^2Y 0.45), and c) lipid extracts ^{1}H NMR spectra of controls (▲, n=15) *vs.* pre-diagnosis GDM (△, n=14) (R^2X 0.37, R^2Y 0.66). Bet: betaine, FA: fatty acids, FC, EC: free and esterified cholesterol, glyc: glyceryl, Lac: lactate, Pyr: pyruvate, TMAO: trimethylamine N-oxide, 1,5-AG: 1,5-anhydroglucitol, Un 3: unknown 3 (δ 5.91). 3-letter codes used for amino acids.

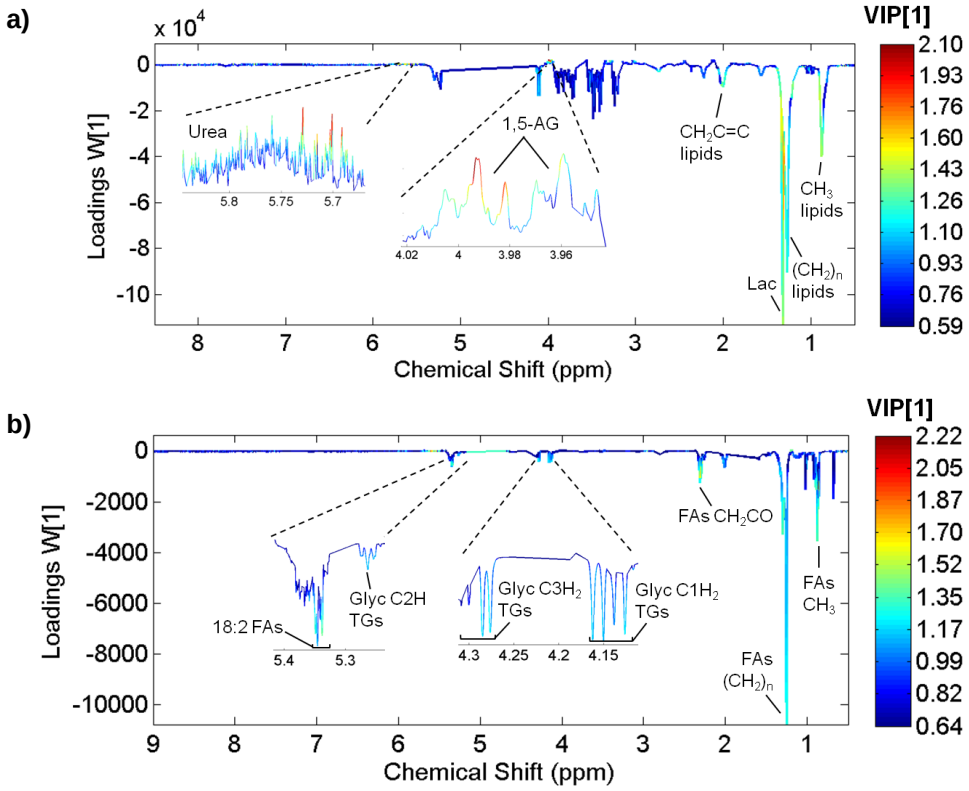


Figure 7.4. PLS-DA loading plots of a) CPMG ^{1}H NMR spectra of plasma for controls n=49 *vs.* pre-diagnosis GDM n=32, and b) ^{1}H NMR spectra of lipid extracts for controls n=15 *vs.* pre-diagnosis GDM n=15. 1,5-AG: 1,5-anhydroglucitol, Glyc: glyceryl.

Specific information on lipid changes preceding GDM was obtained considering the plasma diffusion-edited spectra and lipid extract spectra (Figure 7.3b,c and Figure 7.4b), which showed increases in several lipid resonances (cholesterol, HDL and LDL+VLDL). Out of the variations listed above, Bonferroni correction of p-values identified the increase in the $C18H_3$ cholesterol resonance (measured in diffusion-edited spectra) as the only statistically relevant change (f in Table 7.2, symbol above the $p<1.7\times10^{-3}$ threshold in Figure 7.3b, right), in pre-diagnosis GDM. STOCSY results (not shown) determined that this cholesterol resonance originates from free and/or esterified forms of HDL and LDL+VLDL cholesterol. The PLS-DA model obtained for lipid extracts (Figure 7.3c, Table 7.1) indicated additional small increases in FA and TG glyceryl resonances, along with a decrease in unknown 3 (δ 5.91) (Table 7.3).

Notably, when recalculating the CPMG PLS-DA model using the integral values found to change significantly ($p < 0.05$), instead of the full resolution (variable selected) spectra, no robust PLS-DA group separation was obtained, which thus indicates that spectral detail (often not expressible by discrete peak integrals) seems to hold enhanced classification power for pre-diagnosis GDM, compared to a limited set of peak integrals, with residual univariate statistical relevance.

Table 7.2. List of metabolites/resonances selected in the NMR spectra of blood plasma as important for discrimination of pre- and post-diagnosis GDM from controls. [a] Chemical shifts of signals used for integration; s: singlet, d: doublet, dt: doublet of triplets, q: quartet, m: multiplet, br: broad. [b] E.S.: effect size determined as described in (Berben et al., 2012); values in brackets correspond to high uncertainties. [c] 95% significance level (p-value < 0.05); only values with statistical relevance are specified. [d] Integrals measured in the CPMG spectra and [e] in the diffusion-edited spectra. [f] Alterations remaining significant after Bonferroni correction, with cut-off p-value of 1.7×10^{-3} (0.05 divided by 29 resonances) and 1.3×10^{-3} (0.05 divided by 38 resonances) for pre- and post-diagnosis GDM, respectively. † Alterations with a possible contribution of normal gestational age evolution.

Compound	Chemical shift (ppm) [a]	Pre-diagn. GDM (n=32) *vs.* controls (n=49)		Post-diagn. GDM (n=12) *vs.* controls (n=49)	
		E.S. [b]	p-value [c]	E.S. [b]	p-value [c]
Low M_w compounds [d]					
Valine	1.03 (d)	↑ (0.38 ± 0.46)	2.6×10^{-2}	↓	4.5×10^{-2}
Alanine	1.47 (d)			↓ -0.65 ± 0.64	5.0×10^{-2}
Pyruvate	2.36 (s)	↑ 0.49 ± 0.46	4.5×10^{-2}	↓ -0.74 ± 0.65	4.4×10^{-3}
Glutamine	2.44 (m)	↓	>0.05	↓ -1.27 ± 0.67	2.0×10^{-5} f
Citrate	2.52 (d)			↓	>0.05
Creatine	3.03 (s)	↓	>0.05	↓	>0.05
Creatinine	3.04 (s)			↓ -1.69 ± 0.70	3.5×10^{-9} f
Dimethylsulfone	3.14 (s)	↓	>0.05	↓	>0.05
TMAO	3.27 (s)	↓	>0.05	↓ -0.86 ± 0.65	4.8×10^{-3}
Betaine	3.29 (s)	↓	>0.05	↓ -1.40 ± 0.68	3.5×10^{-8} f
Proline	3.34 (dt)	↓ -0.53 ± 0.46	2.1×10^{-2}	↓ -1.33 ± 0.68	2.0×10^{-5} f
Methanol	3.36 (s)			↓ -1.18 ± 0.67	9.4×10^{-6} f
Glycine	3.55 (s)			↓	4.6×10^{-3}
1,5-anhydroglucitol	3.95-4.00 (dd)	↓ -0.48 ± 0.46	3.5×10^{-2}		
Lactate	4.11 (q)	↑	>0.05	↓	>0.05
Glucose	5.23 (d)	↑	>0.05	↓ -0.96 ± 0.66	1.3×10^{-3} f
Urea	5.77 (br)	↓ -0.58 ± 0.46	2.1×10^{-2}		
Unknown 1	7.15-7.35 (br)			↑ 1.62 ± 0.70	4.6×10^{-5} f
High M_w compounds [e]					
C18H cholesterol	0.67 (br)	↑ 1.23 ± 0.49	1.0×10^{-6} †	↑ 1.35 ± 0.68	2.2×10^{-5} †
CH_3 lip. HDL	0.79-0.85 (br)	↑ 0.60 ± 0.46	8.5×10^{-3}	↑ 1.47 ± 0.68	1.3×10^{-6} f
CH_3 lip. LDL+VLDL	0.85-0.91 (br)	↑ 0.82 ± 0.47	4.2×10^{-3}	↑ 2.54 ± 0.78	1.1×10^{-5} f
$(CH_2)_n$ lip. HDL	1.18-1.25 (br)	↑ 0.60 ± 0.46	8.8×10^{-3}	↑ 1.38 ± 0.68	4.0×10^{-5} †
$(CH_2)_n$ lip. LDL+VLDL	1.25-1.37 (br)	↑ 0.69 ± 0.46	3.1×10^{-2}	↑ 1.87 ± 0.71	9.7×10^{-4} f
$C\underline{H_2}CH_2CO$ lip.	1.45-1.62 (br)	↑ 0.56 ± 0.46	2.5×10^{-2}	↑ 1.50 ± 0.68	4.0×10^{-5} f
$C\underline{H_2}CH$= lip.	1.90-2.02 (br)	↑ 0.65 ± 0.46	9.2×10^{-3}	↑ 2.37 ± 0.76	4.8×10^{-9} †
CH_2CO lip.	2.17-2.26 (br)	↑	>0.05	↑ 1.49 ± 0.68	1.1×10^{-4} f
=CHC$\underline{H_2}$CH= lip.	2.65-2.84 (br)	↑	>0.05	↑ 1.29 ± 0.66	5.6×10^{-4} f
$N(CH_3)_3$ chol. HDL [d, †]	3.19-3.205 (br)			↓ -1.15 ± 0.67	1.9×10^{-3}
$N(CH_3)_3$ chol. LDL+VLDL	3.23-3.26 (br)	↑	>0.05	↑ 2.24 ± 0.75	1.7×10^{-7} f
$C\underline{H_2}N(CH_3)_3$ chol. †	3.62-3.68 (br)			↑ 1.29 ± 0.67	2.1×10^{-5} f
Unknown 2	3.84-3.92 (br)			↑ 1.72 ± 0.70	1.2×10^{-5} f
Glyceryl-C1,3H †	4.02-4.10 (br)			↑ 1.00 ± 0.66	2.3×10^{-2}
Glyceryl-C1,3H'	4.21-4.32 (br)			↑ 1.88 ± 0.71	2.8×10^{-7} f
HC=CH HDL	5.24-5.28 (br)	↑ 0.54 ± 0.46	2.4×10^{-2}	↑ 1.71 ± 0.70	3.8×10^{-5} †
HC=CH LDL+VLDL	5.28-5.37 (br)	↑ 0.63 ± 0.46	1.0×10^{-2}	↑ 1.55 ± 0.69	1.7×10^{-6} †
$(CH_2)_n/CH_3$ lip. HDL		↑	>0.05	↑	>0.05
$HC=CH/CH_3$ lip. HDL		↑	>0.05	↑ 1.30 ± 0.67	1.2×10^{-3} f
$(CH_2)_n/CH_3$ lip. LDL+VLDL †		↑	>0.05	↑ 0.85 ± 0.65	3.8×10^{-2}
$HC=CH/CH_3$ lip. LDL+VLDL		↑	>0.05	↑	>0.05
CH_3 HDL/CH_3 LDL+VLDL		↓	>0.05	↓ -1.30 ± 0.67	2.4×10^{-3}

Post-diagnosis GDM vs. controls

Upon GDM diagnosis, separation from controls was clearer in PLS-DA and models were good for both CPMG and diffusion-edited spectra of plasma (Figure 7.5a,b), as well as for lipid extracts (Figure 7.5c), with or without variable selection (Table 7.1). The larger impact of post-diagnosis GDM on plasma composition, compared to that of the pre-diagnosis GDM group, is illustrated in the loading plots (Figure 7.6) and in a longer list of significant metabolite variations, both in plasma (Table 7.2) and lipid extracts (Table 7.3), including upon Bonferroni correction.

Table 7.3. List of metabolites/resonances selected in the NMR spectra of plasma lipid extracts as important for discrimination of pre- and post-diagnosis GDM from controls. [a,b,c] as defined in Table 7.2 caption. [d] Alterations remaining significant after Bonferroni correction, cut-off p-value of 3.1×10^{-3} (0.05 divided by 16 resonances) and 2.8×10^{-3} (0.05 divided by 18 resonances) for pre- and post-diagnosis GDM, respectively.

Compound	Chemical shift (ppm) [a]	Pre-diagn. GDM (n=14) *vs.* controls (n=15)		Post-diagn. GDM (n=12) *vs.* controls (n=15)	
		E.S. [b]	p-value [c]	E.S. [b]	p-value [c]
Fatty acids (FAs)					
CH_3	0.89 (t)	↑ 0.69 ± 0.53	1.2×10^{-2}	↑ 1.25 ± 0.59 [d]	8.5×10^{-5} [d]
$(CH_2)_n$	1.20-1.27 (br)	↑ 0.62 ± 0.53	2.4×10^{-2}	↑ 0.94 ± 0.57	1.6×10^{-3} [d]
$=CHCH_2(C\underline{H}_2)_n$	1.27-1.40 (br)	↑ 0.69 ± 0.53	1.1×10^{-2}	↑ 1.21 ± 0.58	4.4×10^{-4} [d]
$C\underline{H}_2CH_2CH=$	1.93-2.02 (m)	↑ 0.55 ± 0.52	4.1×10^{-2}	↑ 1.21 ± 0.58	1.5×10^{-4} [d]
$C\underline{H}_2CH=$ 18:2 FA	2.05 (q)			↑ 0.68 ± 0.55	2.4×10^{-2}
CH_2CO	2.28-2.34 (m)	↑ 0.78 ± 0.53	4.8×10^{-3}	↑ 1.26 ± 0.59	9.7×10^{-5} [d]
$=CHC\underline{H}_2CH=$ 18:2 FA	2.77 (t)			↑ 0.62 ± 0.55	4.0×10^{-2}
HC=CH 18:2 FA	5.29-5.35 (m)	↑	>0.05	↑ 0.98 ± 0.57	5.3×10^{-3}
HC=CH 20:4 FA	5.35-5.43 (m)	↑	>0.05	↑ 0.74 ± 0.55	5.7×10^{-3}
Free cholesterol (FC) and esterified cholesterol (EC)					
$C18\underline{H}_3$ FC & EC	0.68 (s)	↑	>0.05	↑ 0.73 ± 0.55	2.0×10^{-2}
$C26\underline{H}_3$ FC & EC	0.86 (d)	↑	>0.05	↑ 0.72 ± 0.55	2.1×10^{-2}
$C19\underline{H}_3$ FC	1.01 (s)	↑	>0.05	↑ 0.76 ± 0.56	1.5×10^{-2}
$C19\underline{H}_3$ EC	1.02 (s)	↑	>0.05	↑ 0.69 ± 0.55	2.7×10^{-2}
$C4\underline{H}_2$ EC	2.26 (t)	↑	>0.05	↑ 0.60 ± 0.55	4.8×10^{-2}
C3H EC	4.61 (m)	↑	>0.05	↑ (0.42 ± 0.52)	>0.05
Triglycerides (TGs)					
Glyceryl-C1,3H	4.14 (dd)	↑ 0.56 ± 0.53	3.6×10^{-2}	↑ 1.12 ± 0.58	7.2×10^{-5} [d]
Glyceryl-C1,3H'	4.29 (dd)	↑ 0.61 ± 0.53	4.2×10^{-2}	↑ 1.07 ± 0.57	5.1×10^{-4} [d]
Unknown 3	5.91 (br)	↓ -0.69 ± 0.53	1.2×10^{-2}	↓ -0.73 ± 0.55	1.0×10^{-2}

Variations comprise a general decrease in several low M_w compounds (with the exception of unknown 1, at δ 7.15-7.35, seen to increase) and marked increases in lipids (Figure 7.5). However, these changes may have some contribution from the gestational age difference between controls (average 17 g.w.) and post-diagnosis GDM (average 26 g.w.), so that such dependence needs to be quantified in order for a trustworthy profile of GDM alone to be obtained.

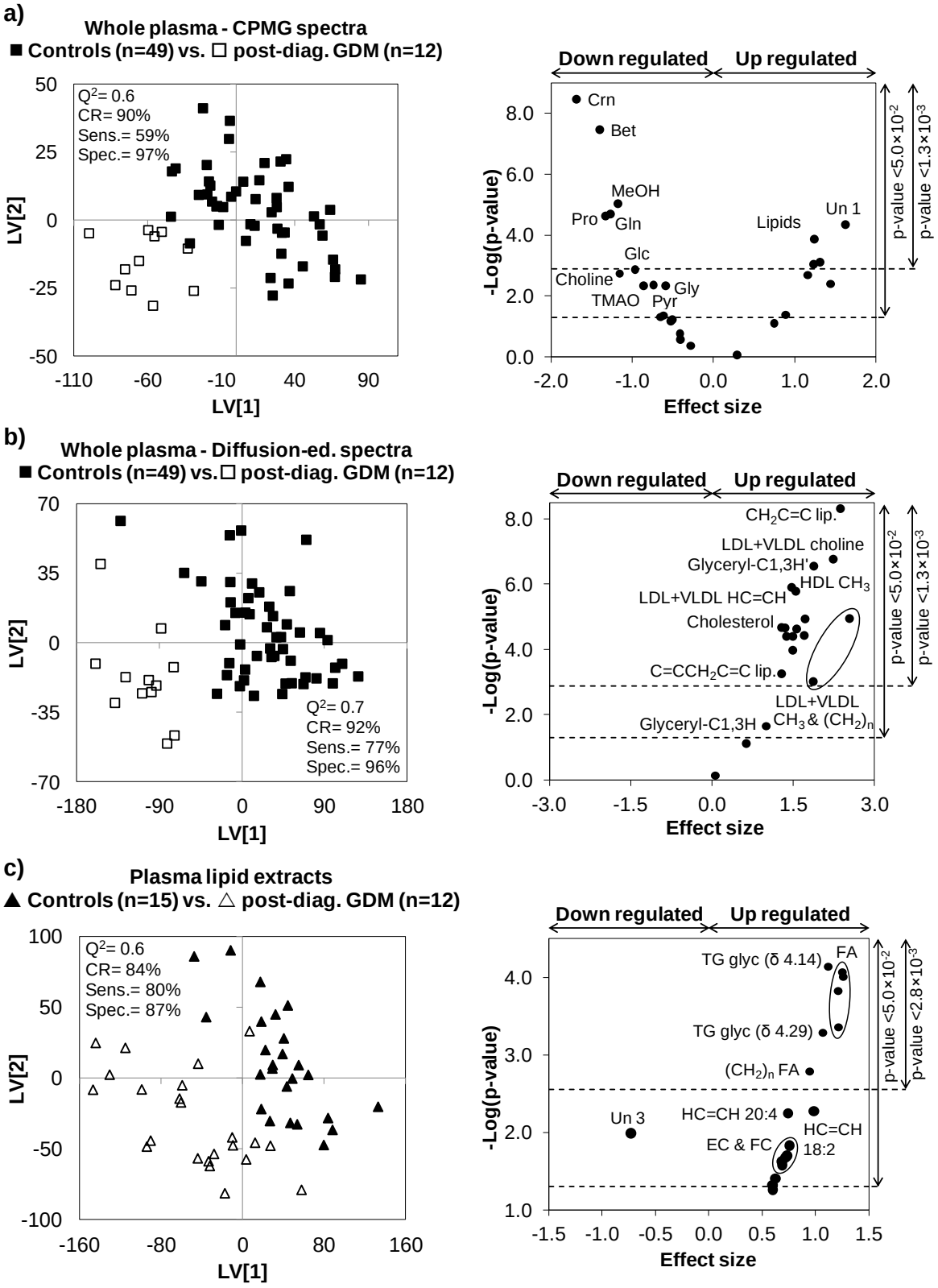


Figure 7.5. PLS-DA scores scatter plots and Volcano plots (effect size vs. – Log (p-value)) of a) plasma CPMG ^{1}H NMR spectra of controls (■, n=49) *vs.* post-diagnosis GDM (□, n=12) (R^2X 0.35, R^2Y 0.69), b) plasma diffusion-edited ^{1}H NMR spectra of controls (■, n=49) *vs.* post-diagnosis GDM (□, n=12) (R^2X 0.47, R^2Y 0.73), and c) lipid extracts ^{1}H NMR spectra of controls (▲, n=15) *vs.* post-diagnosis GDM (△, n=12) (R^2X 0.28, R^2Y 0.72). Bet: betaine, Crn: creatinine, EC: esterified cholesterol, FC: free cholesterol, FA: fatty acids, glyc: glyceryl, Glc: glucose, MeOH: methanol, Pyr: pyruvate, TG: triglycerides, TMAO: trimethylamine N-oxide, Un 1: unknown 1 (δ 7.15-7.35), Un 3: unknown 3 (δ 5.91). 3-letter codes used for amino acids.

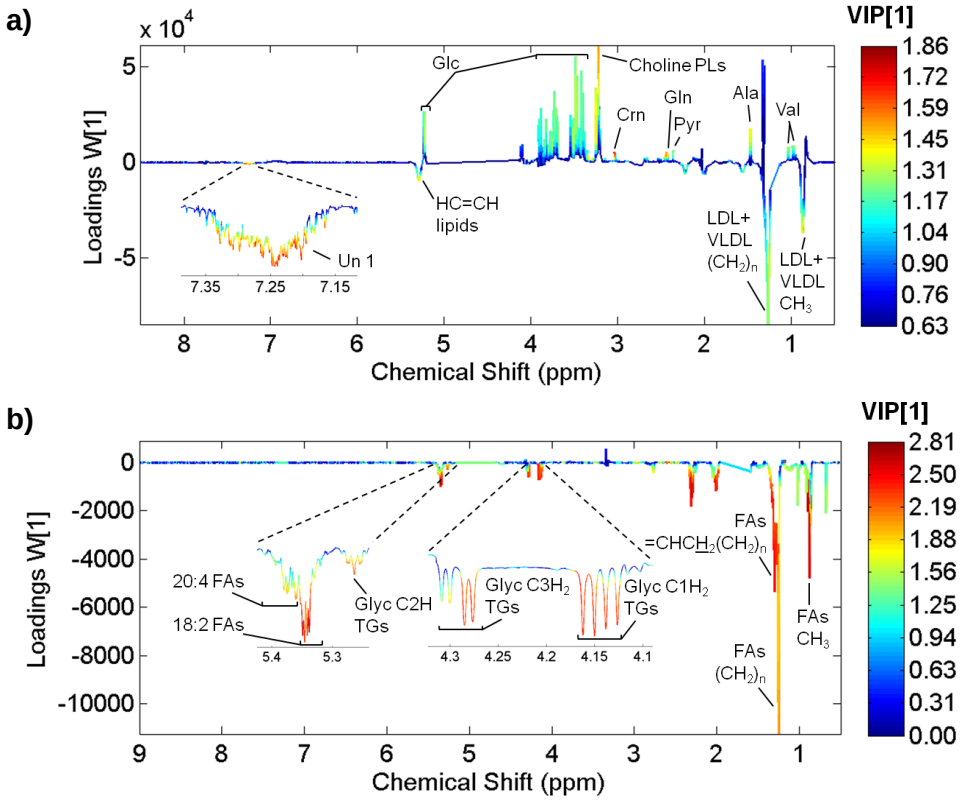


Figure 7.6. PLS-DA loading plots of a) CPMG ^{1}H NMR spectra of plasma for controls n=49 *vs.* post-diagnosis GDM n=12, and b) ^{1}H NMR spectra of lipid extracts for controls n=15 *vs.* post-diagnosis GDM n=12. Crn: creatinine, Glc: glucose, Glyc: glyceryl, Lac: lactate, Pyr: pyruvate, Un 1: unknown 1 (δ 7.15-7.35). 3-Letter codes used for amino acids.

Therefore, a PLS regression was performed between gestational age and peak integrals (with $p < 0.05$), including lipid ratios HDL/(LDL+VLDL) and average FA chain length ($(CH_2)_n/CH_3$) and unsaturation degree ($HC=CH/CH_3$), since these parameters have previously been found to change across pregnancy (Pinto et al., 2014b). The resulting low coefficient of determination (R^2 0.56) (Figure 7.7a) indicated that the set of resonances/ratios correlate poorly with gestational age, in the range considered (16-27 g.w.), thus rendering this as a low-impact variable. For lipid extracts composition, no dependence on gestational age was observed (R^2 0.29, results not shown). The plasma resonances/ratios having a higher time-dependence, across 16-27 g.w., were identified through their b-coefficient values and are situated above the limit of $|\mathbf{b}/\mathbf{b}_{cvSE}| = 1$ (Figure 7.7b). Among these, a subset of 10 resonances/ratios variations had indeed been observed between 2nd and 3rd trimesters of healthy pregnancies(Pinto et al., 2014b) (* in Figure 4b). However, some of those resonances (namely, $C18H_3$ cholesterol; HDL CH_3; LDL+VLDL CH_3, $(CH_2)_n$ and $N(CH_3)_3$ choline; and $CH_2C=C$ of FAs) were also seen to change in the matched group of pre-diagnosis GDM (Table 7.2), which thus shows that their variation, is at least in part, disease dependent. The remaining 4 resonances/ratios (two choline resonances at δ 3.19-3.21 (HDL) and

δ 3.62-3.68; glyceryl-C1,3H resonance at δ 4.02-4.10 and average FA chain length of LDL+LVDL), not noted to change in the matched group, may indeed have a contribution from normal gestational age evolution († Figure 7.7b, Table 7.2), and hence will not be considered as part of the GDM profile. Therefore, we interpret variations in the remaining resonances listed in Table 7.2 as part of the post-diagnosis GDM fingerprint, including those seen to vary due to pregnancy evolution from 16 to 27 g.w. (with $|\mathbf{b}/\mathbf{b}_{cvSE}| \geq 1$ in Figure 7.7b), but as a result of the disease, specifically: betaine, Ala, TMAO, methanol, Pro, creatinine and unknowns 1 (δ 7.15-7.35) and 2 (δ 3.84-3.92).

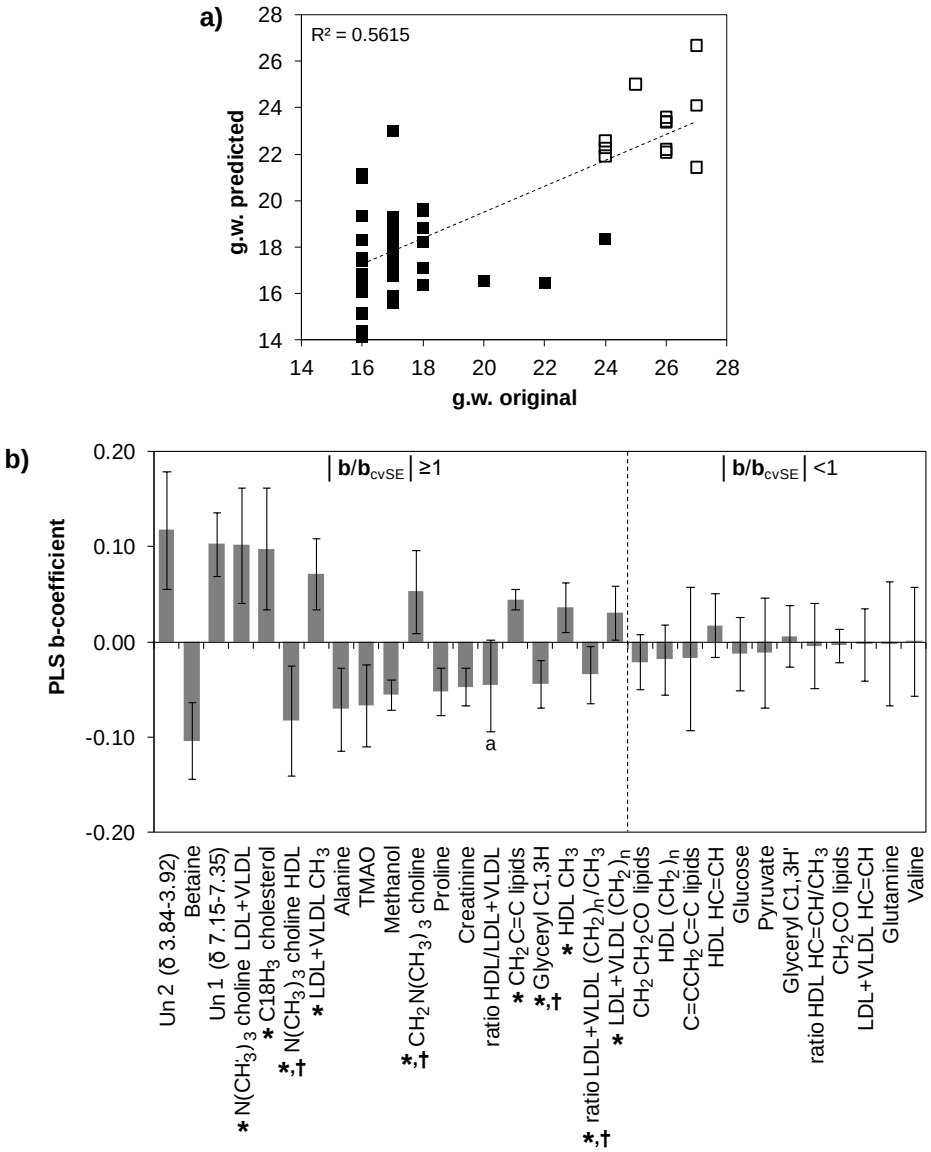


Figure 7.7. PLS regression (LV=2) a) scatter plot of original *vs.* predicted gestational weeks and b) b-coefficients of varying metabolites/resonances, for post-diagnosis GDM compared to controls. [a] ratio with $|b/b_{cvSE}|<1$. * Metabolite variations observed from 2nd to 3rd trimesters of healthy pregnancies (Pinto et al., 2014b), † resonances/ratios with possible contributions from normal gestational age evolution.

Finally, PLS-DA recalculation using the integrals of relevant variations only ($p < 0.05$) (26 resonances altogether, Table 7.2) provided a model with comparable robustness (Q^2 0.6, sens 68%, spec 95%) to that of the model obtained with full resolution variable selected spectra (Table 7.1), thus indicating that a fingerprint composed of such a set of 26 peaks seems sufficiently strong to correctly classify diagnosed GDM samples.

7.2. Proposed metabolic interpretation of plasma changes

This work shows that pre-diagnosis GDM stages (2-21 g.w. prior to diagnosis) involve early metabolic variations (blue arrows in Figure 7.8), detectable in both maternal plasma and corresponding lipid extracts and showing good predictive power when measured through part of the full resolution NMR spectrum, irrespectively of low individual metabolite (univariate) statistical relevance. The noted increase in free and/or esterified cholesterol in plasma was accompanied by smaller increases in plasma HDL, LDL+VLDL lipoproteins and, viewed through lipid extracts, FA (e.g., 18:2 and 20:4) and TG, thus evidencing early disturbances in lipid metabolism (Figure 7.8). Lipid disturbances in pre-diagnostic GDM have been previously suggested through increased TG in serum (dos Santos-Weiss et al., 2013), although other studies have noted the absence of significant variations in total cholesterol, HDL, LDL and TG in plasma (Emet et al., 2013) and serum (Bentley-Lewis et al., 2015) (the latter study reporting, however, a small TG increase). The metabolite profile of pre-diagnosis GDM also comprised a small increase in glucose (possibly related to the observed lower levels of 1,5-anhydroglucitol (Christensen and Williams, 2009)) and some glycolytic and Krebs cycle metabolites (pyruvate, lactate), observed for the first time to our knowledge in pre-diagnosis GDM. Additionally, the particular pattern of amino acid variations (Gln, Pro and Val) differs from a recent report (Bentley-Lewis et al., 2015) of alterations in Ala, Glu and Ser. Other observed changes may stem from deregulation of gut microflora (trimethylamine (TMA)/TMAO, dimethyl sulfone and methanol), homocysteine-methionine conversion (betaine) and urea cycle (urea and creatine) (Figure 7.8). Betaine and TMAO decreases match those observed before in a smaller cohort (Diaz and Pinto et al., 2011), the latter metabolite having, in a separate study (Wang et al., 2011), been identified as arising from gut microflora. Indeed, a contribution from gut microbiota to the development of low-grade inflammation has been reported for type 2 diabetes (Everard and Cani, 2013). A recent report detected a

decrease in creatinine in pre-diagnosis GDM (Bentley-Lewis et al., 2015) but, to our knowledge, no changes in creatine or urea have been observed before.

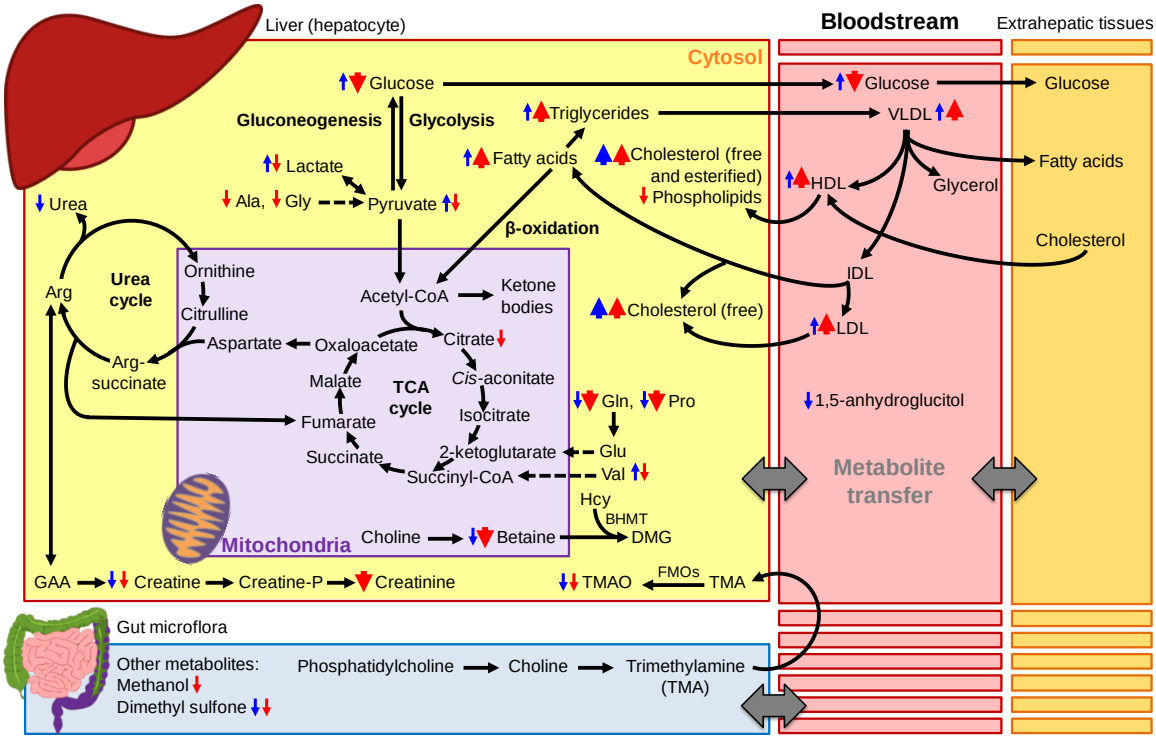


Figure 7.8. Schematic representation of the metabolic pathways possibly affected in pre- (blue arrows) and post-diagnosis GDM (red arrows). Thicker arrows indicate alterations remaining significant after Bonferroni correction. BHMT: betaine homocystein methyltransferase, DMG: dimethylglycine, FMOs: flavin-containing monooxygenases, GAA: guanidinoacetate, Hcy: homocysteine. 3-letter codes used for amino acids.

Many of the changes observed in the pre-diagnosis GDM group were intensified at the time of diagnosis (red arrows in Figure 7.8), whereas new variations were noted for Ala, Gly, citrate, creatinine and methanol. Since control and post-diagnosis GDM groups presented a small difference in gestational age, evaluation of metabolite variations as to their dependence on this parameter revealed that some PL and glyceryl resonances may have a stronger influence of normal gestational age evolution so these were, hence, left out of a putative GDM profile; the same applied to changes seen in average FA chain length of LDL+VLDL lipoproteins. Overall, GDM-related metabolite changes were found to be either independent of gestational age or to have time-dependences (within 16-27 g.w.) determined by the disease. A surprising observation related to the diminished glucose in post-diagnosis GDM subjects, with clear reversed impacts on pyruvate and lactate (Figure 7.8). This contradicts previous reports on GDM groups (Dudzik et al., 2014; Huynh et al., 2014), although a study has

reported no change in glucose (Tarim et al., 2004) and a decrease in this sugar has been observed in umbilical cord blood of GDM-born infants (Dani et al., 2014). A glucose decrease may be explained by the fact that subjects were informed of GDM diagnosis approximately two weeks prior to sample collection having, probably, privately adopted lower sugar diets prior to the formal layout of the GDM management scheme. Subsequent studies should therefore involve sample collection no later than the day of diagnosis, as actually carried out in some studies (Dudzik et al., 2014). The impact of glucose reversal seems to affect lactate and pyruvate but other changes, namely on lipids, cholesterol, amino acids (with the exception of Val) are maintained from pre-diagnosis GDM and, in many cases, intensified thus showing that most of the disease effects seem to remain, in spite of the glucose-related adaptation. In fact, the marked increases in lipid moieties are consistent with previous reports of increased TG, FA and cholesterol (Huynh et al., 2014). Regarding the decreases in amino acids which help replenish Krebs cycle substrates (Gln, Pro, Val, Ala and Gly), all have been seen to vary in different studies of GDM (Huynh et al., 2014), albeit somewhat inconsistently. Furthermore, changes in gut microflora metabolites (including TMA conversion to TMAO) are maintained from the early stages of GDM, whereas both creatine and creatinine are decreased (the latter having been seen to change differently in distinct GDM cohorts (Everard and Cani, 2013; Dudzik et al., 2014; Scholtens et al., 2014)), with urea remaining unchanged, contrary to pre-diagnosis GDM.

7.3. Conclusions

This work enabled some considerations to be drawn in relation to possible future applications of NMR metabolomics in clinical GDM management. Firstly, analysis of plasma seems preferable to lipid extracts, for both pre- and post-diagnosis GDM, since classification performance was comparable for both sample types and direct plasma analysis avoids time-consuming extraction protocols, with less than ideal reproducibility. Furthermore, pre-diagnosis GDM recognition should involve the definition of an early biomarker profile as a full resolution subset of the CPMG ^{1}H NMR spectra, rather than of a discrete set of metabolites, as shown by the higher classification power of the former. This concept of a spectral marker accommodates the assumption of metabolite inter-dependence and multivariate nature of the metabolite profile used to recognize the early stages of GDM. On the other hand, for post-diagnosis GDM, a robust profile of 26 resonances (corresponding to 10

metabolites and lipids) may be defined to describe the condition, as viewed by NMR, thus advancing the possibility of using a multi-metabolite biomarker of GDM as a complementary tool in the clinical management of the disease. In the follow-up of this work, it is important that the putative metabolic disturbances hereby suggested to relate to pre- and post-diagnosis GDM may be demonstrated in larger cohorts and by complementary biochemical methods. In addition, issues such as improving inter-laboratory consistency of experimental design and data handling protocols, as well as considering populations of different ethnicity, would greatly contribute to resolving apparent discrepancies between existing studies and understanding possible cohort-dependent responses to GDM.

Chapter 8. Second trimester maternal plasma for diagnosis and prediction of other prenatal disorders

Part of the results presented in this chapter were published in the article below and are here adapted slightly onto the form of a thesis chapter, for the sake of clarity:

Pinto, J.; Maciel, E.; Melo, T. S.; Domingues, M. R. M.; Galhano, E.; Pita, C.; Almeida, M. do C.; Carreira, I. M.; Gil, A. M. Maternal Plasma Phospholipids Are Altered in Trisomy 21 Cases and prior to Preeclampsia and Preterm Outcomes. *Rapid Communications in Mass Spectrometry* **2014**, *28*, 1635–1638.

8.1. Fetal malformations

The work presented in this subchapter aims to explore the potential of 2nd T maternal plasma to detect metabolic deviations due to fetal malformations (FM) using the whole ^{1}H NMR blood metabolome. Samples metadata is shown in Chapter 2 (Section 2.1.2, Table 2.6 and Table 2.7). The results shown here were not subjected to publication.

8.1.1. Analysis of whole plasma by NMR spectroscopy

Figure 8.1 shows the average CPMG ^{1}H NMR spectra of controls and FM samples. The visual comparison of these spectra revealed subtle changes (arrows in Figure 8.1b) mainly in acetone, acetoacetate, citrate and formate. PCA of the CPMG ^{1}H NMR spectra of 2nd T blood plasma of controls *vs.* FM samples (not shown) unveiled no trends and the corresponding PLS-DA model had poor performance (Q^2 0.1, CR 70%, sens. 55%, spec. 79%). In addition, PCA and PLS-DA of the standard and diffusion-edited spectra unveiled no trends.

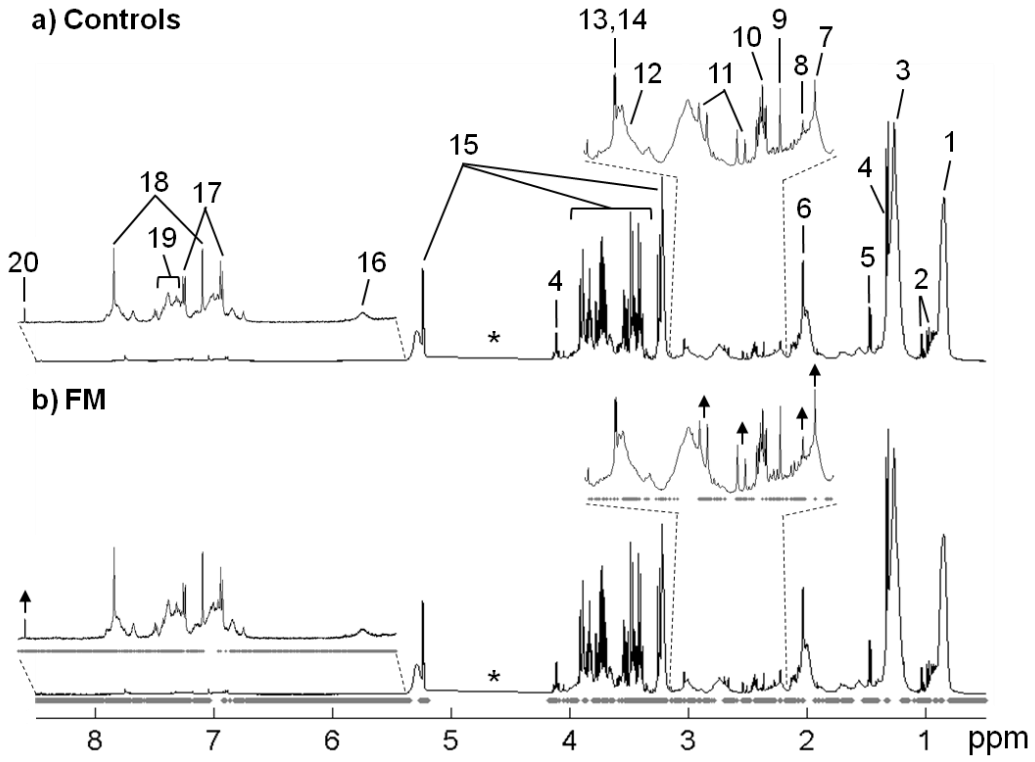


Figure 8.1. Average 500 MHz CPMG ^{1}H NMR spectra of 2nd T plasma of a) controls and b) FM, with indication of spectral variables (data points) selected in the corresponding PLS-DA models (grey dots under spectra). 1- CH_3 lipids, 2- valine, 3- $(CH_2)_n$ lipids, 4- lactate, 5- alanine, 6- N-acetyl glycoproteins, 7- acetone, 8- acetoacetate, 9- pyruvate, 10- glutamine, 11- citrate, 12- albumin-lysyl, 13- creatine, 14- creatinine, 15- glucose, 16- urea, 17- histidine, 18- tyrosine, 19- unknown 1 (δ 7.20-7.39), 20- formate. *: cut off spectral region corresponding to water. Arrows indicate visible alterations.

Variable selection was applied to the CPMG ^{1}H NMR spectra and the subsequent PLS-DA model (Figure 8.2a left) was indeed much improved (Q^2 0.5, CR 81%, sens. 72%, spec. 87%, Figure 8.3a). The grey dots under the CPMG spectrum in Figure 8.1a represent the data points selected as main contributors for FM group separation from controls in Figure 8.2a. Computed metabolite changes (Table 8.1) showed that FM samples carry lower amounts of urea and $N(CH_3)_3$ choline of LDL+VLDL, and increased levels of β-hydroxybutyrate, citrate and acetoacetate, as well as qualitative decrease in methanol and increases in N-acetyl glycoproteins, acetone, albumin-lysyl, pyruvate, CH_3 lipids of LDL+VLDL and lactate, as clearly seen in the Volcano plot (Figure 8.2a right).

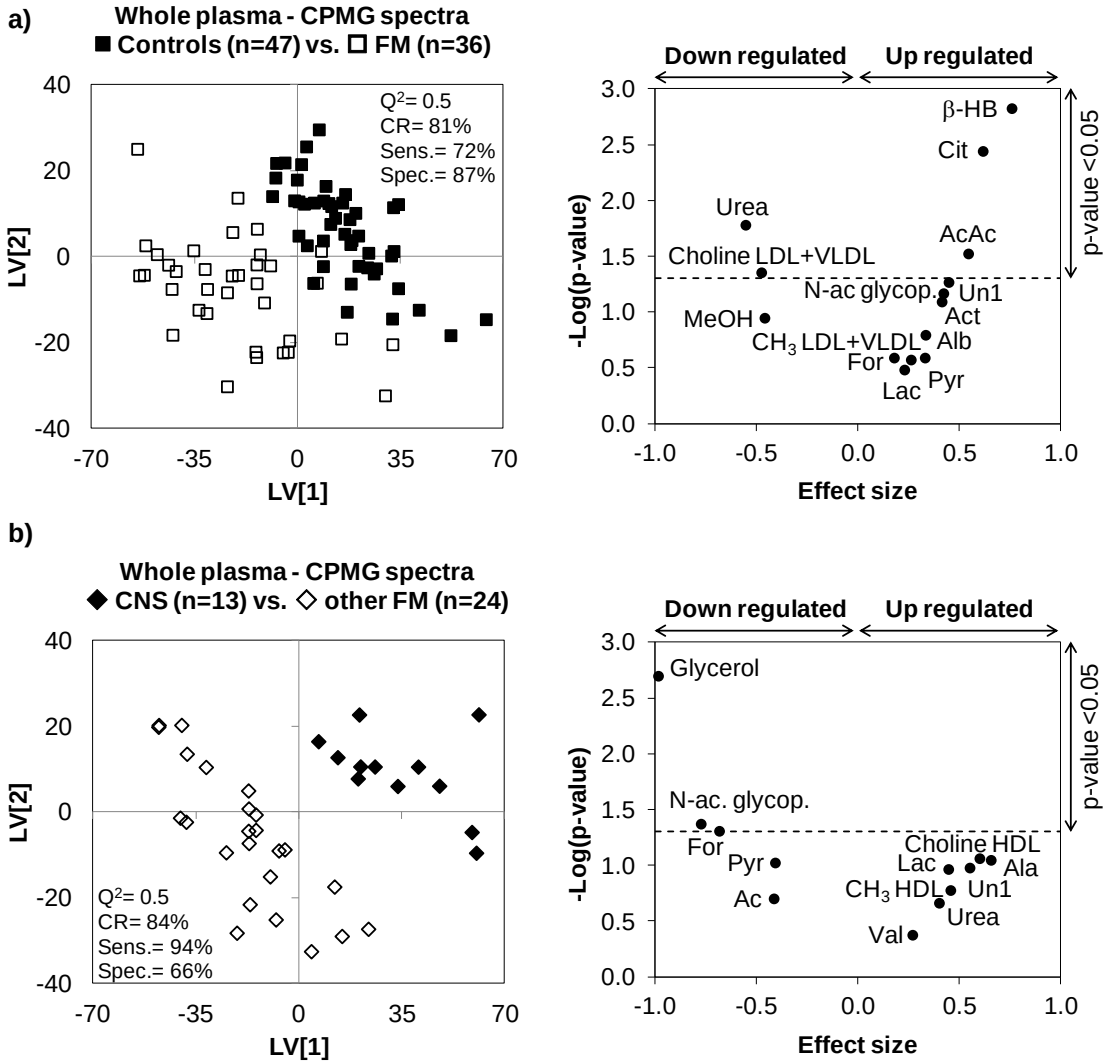
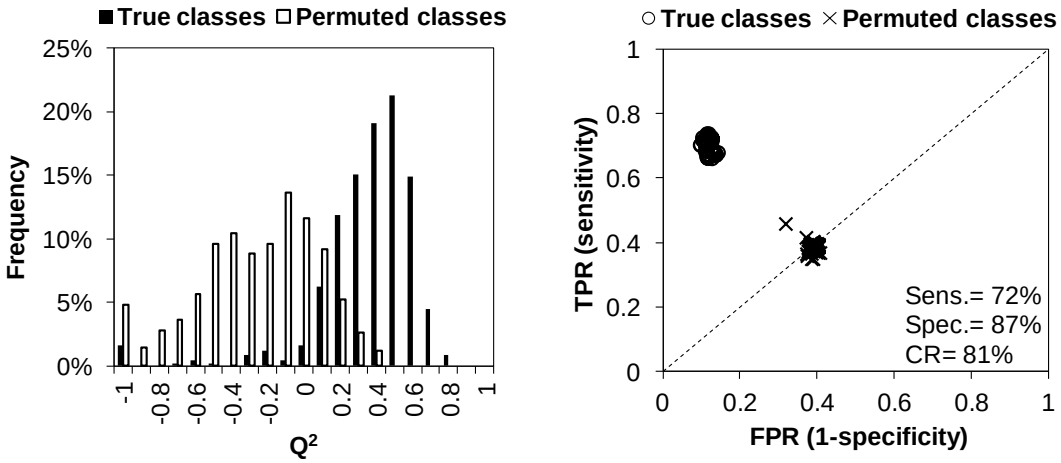


Figure 8.2. PLS-DA scores scatter plots and Volcano plots (effect size *vs.* – Log (p-value)) of plasma CPMG ^{1}H NMR spectra of a) controls (■, n=47) *vs.* FM (□, n=37) (R^2X 0.14, R^2Y 0.77), and b) CNS (◆, n=13) *vs.* other FM (◇, n=24) (R^2X 0.23, R^2Y 0.87). AcAc: acetoacetate, Alb: albumin-lysyl, Act: acetone, cit: citrate, For: formate, N-ac. glycop: N-acetyl glycoproteins, β-HB: β-hydoxybutyrate, Lac: lactate, Pyr: pyruvate, Un1: unknown 1 (δ 7.20-7.39). 3-letter codes used for amino acids.

When comparing CNS cases with other FM cases, the PCA of CPMG ^{1}H NMR spectra revealed no trends (not shown) and the corresponding PLS-DA model had poor performance (Q^2 0.1, CR 62%, sens. 79%, spec. 32%). The PLS-DA model (Figure 8.2b left) obtained after variable selection was much improved (Q^2 0.5, CR 84%, sens. 94%, spec. 66%, Figure 8.3b). The other FM group includes different malformation types (as described in Chapter 2, section 2.1.2, Table 2.7), namely cardiac (n=7), polimalformations (n=6), limbs (n=4), urogenital (n=3) and other types (n=4). Subsequent loadings inspection (not shown) and signal integration unveiled that CNS malformations have a distinct profile compared with other FM comprising changes with statistical relevance ($p < 0.05$) in glycerol (↓) and N-acetyl glycoproteins (↓), tandem with variation tendencies in pyruvate (↓), acetate (↓), CH_3 and choline HDL (↑), valine (↑), alanine (↑) and lactate (↑) (Figure 8.2b right).

a) Controls (n=47) vs. FM (n=36) – CPMG spectra after variable selection



b) CNS (n=13) vs. other FM (n=24) – CPMG spectra after variable selection

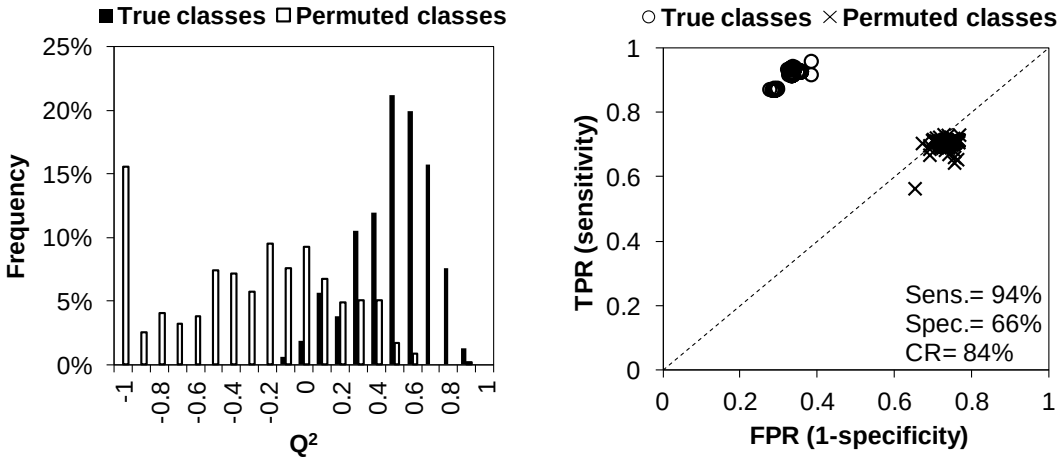


Figure 8.3. Q^2 distributions (left) and ROC plots (right) of true and permuted models obtained for CPMG ^{1}H NMR spectra of whole blood plasma after variable selection of a) controls (n=47) *vs.* FM (n=37) and b) CNS (n=13) *vs.* other FM (n=24).

Table 8.1. List of metabolites/resonances selected in the ^{1}H CPMG NMR spectra of blood plasma as important for discrimination of FM from controls and CNS from other FM. [a] Chemical shifts of signals used for integration; s: singlet, d: doublet, dd: doublet of doublets, q: quartet, br: broad. [b] E.S.: effect size determined as described in (Berben et al., 2012). [c] 95% significance level ($p < 0.05$); only values with statistical relevance are specified.

Compound	Chemical shift (ppm) [a]	FM (n=36) *vs.* controls (n=47)		CNS (n=13) *vs.* other FM (n=24)	
		E.S. [b]	p-value [c]	E.S. [b]	p-value [c]
CH_3 HDL	0.79-0.86 (br)			↑	> 0.05
CH_3 LDL+VLDL	0.86-0.91 (br)	↑	> 0.05		
Valine	1.03 (d)			↑	> 0.05
β-hydroxybutyrate	1.19 (d)	↑ 0.76 ± 0.45	1.50×10^{-3}		
Alanine	1.45 (d)			↑	> 0.05
N-acetyl glycoproteins	2.03 (s)			↓ -0.77 ± 0.70	4.27×10^{-2}
Acetone	2.22 (s)	↑	> 0.05	↓	> 0.05
Acetoacetate	2.27 (s)	↑ 0.55 ± 0.45	3.01×10^{-2}		
Pyruvate	2.36 (s)	↑	> 0.05	↓	> 0.05
Citrate	2.52 (d)	↑ 0.62 ± 0.45	3.63×10^{-3}		
Albumin-lysyl	2.92-3.03 (br)	↑	> 0.05		
$N(CH_3)_3$ cho. HDL	3.18-3.22 (br)			↑	> 0.05
$N(CH_3)_3$ cho. LDL+VLDL	3.22-3.24 (br)	↓ -0.47 ± 0.44	4.44×10^{-2}		
Methanol	3.36 (s)	↓	> 0.05		
Glycerol	3.65 (dd)			↓ -0.98 ± 0.71	2.01×10^{-3}
Lactate	4.11 (q)	↑	> 0.05	↑	> 0.05
Urea	5.77 (br)	↓ -0.55 ± 0.45	1.66×10^{-2}	↑	> 0.05
Unknown	7.20-7.39 (br)	↑	> 0.05	↑	> 0.05
Formate	8.45 (s)	↑	> 0.05	↓	> 0.05

When recalculating the CPMG PLS-DA model using the integral values found to change significantly ($p < 0.05$), instead of the full resolution (variable selected) spectra, no reliable PLS-DA group separation was achieved for both controls *vs.* FM and CNS *vs.* other FM. This finding indicates that spectral detail (full spectral resolution) needs to be taken into account in order to achieve enhanced classification power, compared to a limited set of peak integrals, as previously observed for other disorders.

8.1.2. Proposed metabolic interpretation of plasma changes

The metabolic alterations found in 2nd T maternal plasma of pregnant women carrying malformed fetuses were consistent with previously observed changes in a lower cohort (25 controls and 26 FM) (Diaz and Pinto et al., 2011), namely in increased levels of CH_3 of lipids, ketone bodies (β-hydroxybutyrate, acetone and acetoacetate) and a decreased tendency for methanol. The increased levels of ketone bodies may reflect an increased oxidation of fatty acids in liver and ketone bodies transfer through blood for conversion to acetyl-CoA and oxidation in the TCA cycle for energy production in extra-hepatic tissues. Alterations in ketone bodies consistent with lipid

oxidation were also found for 2nd T maternal urine of FM cases (Diaz et al., 2013b), along with changes in carnitine and other metabolites in 2nd T amniotic fluid (Graça et al., 2012b). Interestingly in this work, two resonances from LDL+VLDL were seen to change in FM cases but in opposite directions, while the CH_3 resonance from total fatty acids tended to increase, the $N(CH_3)_3$ resonance from phospholipids showed a decrease tendency. This suggests a depletion of phospholipids in LDL+VLDL particles, seen here for the first time to our knowledge. For congenital heart defects, lower levels of phospholipids (including several PC, LPC and SM) have been reported for 1st T maternal deproteinized serum by LC-MS/MS (Bahado-Singh et al., 2014). Regarding the decreasing tendency of methanol, consistent with our previous report (Diaz and Pinto et al., 2011), a possible explanation may be related with alterations in gut microflora (Komarova et al., 2014).

Novel alterations in 2nd T maternal plasma of FM cases comprised the increase in pyruvate, lactate, formate, citrate, albumin, and the unknown resonance at δ 7.20-7.39, as well as the decrease in urea levels. The increasing tendency of pyruvate, lactate and formate suggests enhanced glycolysis consistent with the glucose depletion observed in amniotic fluid (Graça et al., 2010). In addition, increased levels of citrate seem to reflect its underuse in TCA cycle. Alterations in pyruvate and lactate levels in 1st and 2nd T maternal sera were observed before for NTD cases but in the opposite direction suggesting a reduction in glycolysis activity (Zheng et al., 2011). An increased tendency for albumin may suggest increased protein biosynthesis, which has also been previously found in amniotic fluid (Graça et al., 2010). Relevant changes were also observed for urea levels (lower), possibly reflecting alterations in urea cycle and/or in renal function. Alterations in renal function were found before in 2nd T amniotic fluid of FM cases (Graça et al., 2010), whereas for 2nd T maternal urine no changes were reported .

CNS malformations showed a distinct plasma metabolic profile when compared to other FM types. Statistically significant changes comprise the decrease in glycerol and N-acetyl glycoproteins, while only qualitative changes were observed for HDL (CH_3 and $N(CH_3)_3$ choline resonances), valine, alanine, acetone, pyruvate, lactate, urea, formate and unknown resonance at δ 7.20-7.39. Changes in valine and lipid profile in CNS, compared with the remaining malformations, are consistent with our previous report (Diaz and Pinto et al., 2011). Novel alterations include the lower levels of glycerol, possibly reflecting enhanced glycerolipid metabolism with glycerol utilization to form TG and other glycerolipids. Other significant change found in this work is the decrease in N-acetyl glycoproteins metabolism suggesting a decrease in synthesis of a number of proteins (α1-acid glycoprotein, α1-antitrypsin and haptoglobulin (Torri et al., 1999)), not

previously observed for CNS malformations. Furthermore, reversed variations were seen for acetone, pyruvate, formate and urea in CNS vs. other FM compared with controls vs. FM, suggesting a possible lower contribution of CNS cases for ketone body synthesis, enhanced glycolysis and alterations in urea cycle or renal function.

8.2. Pre-premature rupture of membranes (pre-PROM)

In this subchapter, the possibility of predicting PROM (here named pre-PROM), occurring after the 37th g.w., through 2nd T maternal plasma was attempted by comparing samples collected 19-24 g.w. before the PROM diagnosis with controls. A more complete description of samples is shown in Chapter 2 (Section 2.1.2., Table 2.6). The results shown here were not published.

8.2.1. Analysis of whole plasma by NMR spectroscopy

Figure 8.4 shows the average CPMG ^{1}H NMR spectra of controls and pre-diagnosis PROM (pre-PROM) where no visible changes are noted for pre-PROM group. A strong impact of pre-PROM on maternal plasma metabolome is not expected since this condition is an integral part of the normal parturition process at term without posing significant risks for mother and fetus.

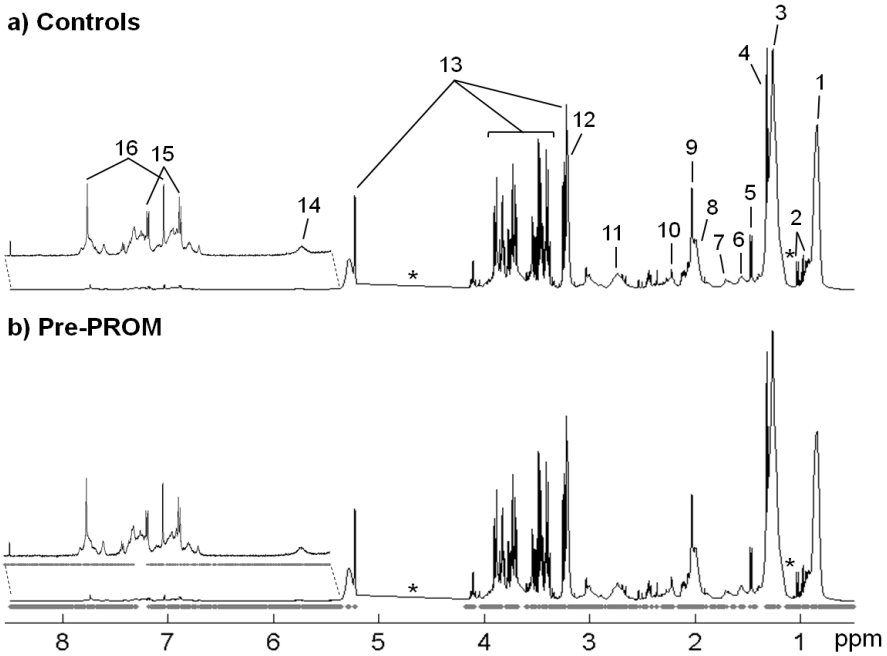


Figure 8.4. Average 500 MHz CPMG ^{1}H NMR spectra of 2nd T plasma of a) controls and b) pre-PROM, with indication of spectral variables (data points) selected in the corresponding PLS-DA models (grey dots under spectra). 1- CH_3 lipids, 2- valine, 3- $(CH_2)_n$ lipids, 4- lactate, 5- alanine, 6- $C\underline{H_2}CH_2CO$ lipids, 7- $C\underline{H_2}CH_2CH=$ lipids, 8- $C\underline{H_2}CH=$ lipids, 9- N-acetyl glycoproteins, 10- acetone superimposed on CH_2CO lipids, 11- $=CHC\underline{H_2}CH=$ lipids, 12- $N(CH_3)_3$ choline of PL, 13- glucose, 14- urea, 15- histidine, 16- tyrosine. *: cut off spectral region corresponding to water.

Multivariate analysis of standard and diffusion-edited spectra produced models with no trends between controls and pre-PROM. PLS-DA of full resolution CPMG ^{1}H NMR spectra produced a poor model with low predictive power and robustness (Q^2 0.1, CR 71%, sens. 58%, spec. 81%). After variable selection, an improved model was obtained (Figure 8.5) (Q^2 0.5, CR 82%, sens. 73%, spec. 89%) with a separation trend between the two groups although with 8 pre-PROM samples overlapped with controls, thus justifying the low sensitivity. Loadings inspection and spectral integration unveiled four statistical significant alterations ($p < 0.05$) and ten qualitative changes in plasma resonances of pre-PROM compared to controls, as listed in Table 8.2.

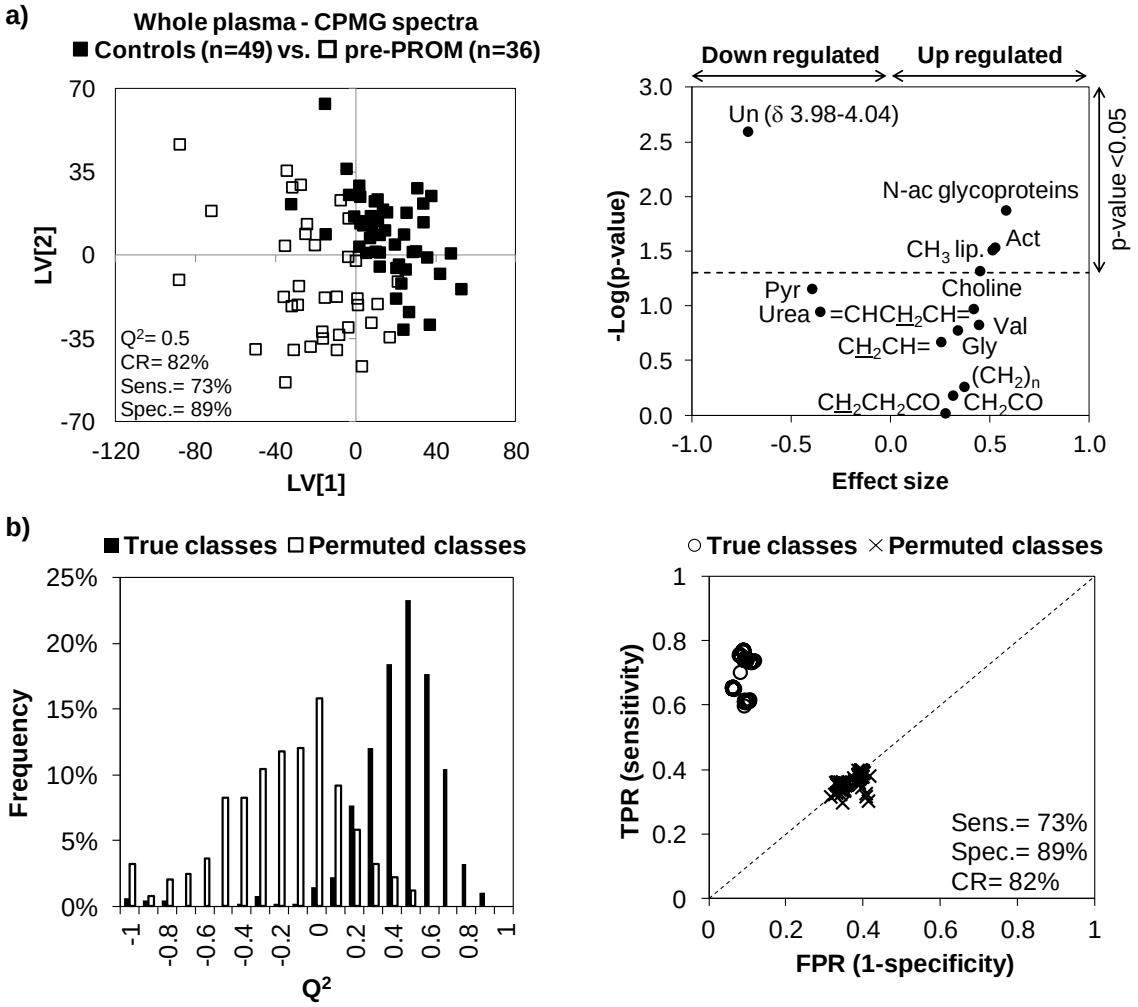


Figure 8.5. a) PLS-DA scores scatter plot (left) and Volcano plots (effect size *vs.* – Log (p-value)) (right) and b) Q^2 distributions (left) and ROC plots (right) of true and permuted models obtained for plasma CPMG ^{1}H NMR spectra of a) controls (■, n=47) *vs.* pre-PROM (□, n=37) (R^2X 0.26, R^2Y 0.62). Act: acetone, Pyr:pyruvate and 3-letter code for amino acids.

Significant alterations included a small increase in CH_3 lipids (total), N-acetyl glycoproteins and acetone, and decrease in an unknown resonance at δ 3.98-4.04

possibly arising from Hα resonances of amino acids. Qualitative changes included several lipidic resonances ($(CH_2)_n$, $C\underline{H}_2CH_2CO$, $CH_2CH=$, CH_2CO, $=CHC\underline{H}_2CH=$ and $N(CH_3)_3$ choline), glycine and urea.

Table 8.2. List of metabolites/resonances selected in the ^{1}H CPMG NMR spectra of blood plasma as important for discrimination of pre-PROM from controls. [a] Chemical shifts of signals used for integration; s: singlet, d: doublet, m: multiplet, br: broad. [b] E.S.: effect size determined as described in (Berben et al., 2012). [c] 95% significance level (p-value < 0.05); only values with statistical relevance are specified.

Compound	Chemical shift (ppm) [a]	Pre-PROM (n=36) *vs.* controls (n=49) E.S. [b]	p-value [c]
CH_3 lipids	0.66-0.91 (br)	↑ 0.51 ± 0.44	3.09×10^{-2}
Valine	1.03 (d)	↑	> 0.05
$(CH_2)_n$ lipids	1.21-1.29 (br)	↑	> 0.05
$C\underline{H}_2CH_2CO$ lipids	1.52-1.62 (br)	↑	> 0.05
$C\underline{H}_2CH=$ lipids	1.93-2.02 (br)	↑	> 0.05
N-acetyl glycoproteins	2.03 (s)	↑ 0.58 ± 0.44	1.34×10^{-2}
Acetone	2.22 (s)	↑ 0.53 ± 0.44	2.93×10^{-2}
CH_2CO lipids	2.18-2.23 (br)	↑	> 0.05
Pyruvate	2.36 (s)	↓	> 0.05
$=CHC\underline{H}_2CH=$ lipids	2.70-2.83 (br)	↑	> 0.05
$N(CH_3)_3$ choline	3.18-3.21 (br)	↑	> 0.05
Glycine	3.55 (s)	↑	> 0.05
Unknown	3.98-4.04 (m)	↓ -0.72 ± 0.44	2.54×10^{-3}
Urea	5.77 (br)	↓	> 0.05

Finally, when recalculating the CPMG PLS-DA model using the integral values found to change significantly ($p < 0.05$), no robust PLS-DA group separation was obtained indicating that the full resolution spectra is important for pre-PROM discrimination. Besides, the altered metabolic profile found for pre-PROM is very weak but any impact is surprising due to the large time between collection and PROM event. Thus, the results found in this work really need confirmation in larger cohorts.

8.2.2. Proposed metabolic interpretation of plasma changes

The impact of pre-PROM on 2nd T maternal plasma metabolome was firstly study considering a lower sample cohort (20 controls and 18 pre-PROM) unveiling only faint alterations in some samples for acetate, glutamine, citrate and albumin (Diaz and Pinto et al., 2011). Considering the larger cohort presented here, none of the previously observed alterations were confirmed. Results obtained included relevant changes in N-acetyl glycoproteins, acetone and unknown resonance at δ 3.98-4.04, along with qualitative changes in several lipidic resonances, valine, pyruvate, glycine and urea. A

higher level of N-acetyl glycoproteins possibly suggests an increased biosynthesis of these proteins seen here for the first time to our knowledge. Indeed, the fetal membrane weakening involved in PROM event has been related with biochemical processes including degradation of membrane's collagen with involvement of proteolytic enzymes (Moore et al., 2006). For 2nd T amniotic fluid, a possible increase of proteins was found for pre-PROM by MIR (Graça et al., 2013), whilst for maternal urine no alterations were observed in NMR metabolic profile. Regarding the significant decrease of the unknown resonance at δ 3.98-4.04 (possible arising from amino acid Hα resonances) and the increase in acetone, no possible metabolic interpretation may be advanced at this stage. Qualitative alterations suggest subtle alterations in lipid metabolism, amino acid catabolism (glycine, valine and pyruvate) and urea cycle. These results showed an apparently specific metabolic signature of 2nd T maternal plasma possible useful for prediction of PROM cases 19-24 g.w. before diagnosis. The complementary use of more sensitive techniques and proteomic approaches could provide more knowledge about the small alterations detected here.

8.3. Pre-diagnostic preterm delivery and preeclampsia

The work presented in this subchapter shows the results obtained for plasma phospholipid profile of pre-diagnostic preterm delivery and preeclampsia by HILIC-LC/MS. Regarding NMR plasma metabolome, no separation trends were obtained for pre-diagnostic preterm delivery (n=16) while for preeclampsia the number of samples (n=4) was insufficient for multivariate analysis. The metabolic alterations obtained for these disorders through phospholipid profiling by HILIC-LC/MS were published as a short communication with reference: Pinto, J.; Maciel, E.; Melo, T. S.; Domingues, M. R. M.; Galhano, E.; Pita, C.; Almeida, M. do C.; Carreira, I. M.; Gil, A. M. Maternal Plasma Phospholipids Are Altered in Trisomy 21 Cases and prior to Preeclampsia and Preterm Outcomes. *Rapid Communications in Mass Spectrometry* **2014**, *28*, 1635–1638. The results concerning the T21 group were described before in Chapter 6 (section 6.2), and here only the results obtained for pre-diagnostic preterm delivery and preeclampsia will be presented. Samples metadata can be found in Chapter 2 (Table 2.6).

Abstract

Rationale: To search for changes in phospholipids in 2nd T maternal plasma, in connection to pre-diagnostic preterm and preeclampsia, aiming at finding predictive

markers of these conditions. Methods: Plasma was collected from pregnant women at 16-19 g.w. defining three groups: controls (n=14), pre-diagnostic preterm delivery (n=5) and pre-diagnostic preeclampsia (n=4). The PL extracts of the pooled samples were analysed by hydrophilic interaction liquid chromatography mass spectrometry (HILIC-LC/MS) and subsequent multivariate and univariate analysis. Results: A measurable impact on maternal PL was noted for pre-diagnostic preterm and preeclampsia. Pre-diagnostic preterm samples comprised changes in unsaturated PC and a PC oxidation product, possibly due to non-enzymatic oxidative processes. PL sensitivity to pre-diagnostic preeclampsia was confirmed by a newly observed LPC changes (common to T21, possibly reflecting a common oxidative mechanism) and specific changes in a PE and an ether-linked PC. Conclusions: Results indicated that maternal PL metabolism is exquisitely sensitive to pre-diagnostic preterm and preeclampsia. The apparent specificity of PL profiles enables future uses for predictive diagnosis to be envisaged, for improved prenatal health management.

Brief state of the art

Lipid changes are known to accompany the onset and evolution of many diseases. In pregnancy, maternal blood lipids naturally undergo extensive changes and their deregulation has been associated with some prenatal disorders. For preeclampsia for instance, changes have been noted in the levels of TG and specific PL, both at diagnosis and at pre-diagnostic stages (10-20 g.w.) (Wiznitzer et al., 2009; Kenny et al., 2010; De Oliveira et al., 2012; Gallos et al., 2013). Furthermore, increased risk of spontaneous preterm birth has been suggested for women with high total cholesterol, LDL-cholesterol and TG (Mudd et al., 2012), whereas altered levels of glycerolipids, PL, sterol lipids and sphingolipids have been found in pregnancies with poor outcomes (stillbirth, preterm, small for gestational age) (Heazell et al., 2012). Preeclampsia and preterm usually pose significant risks in pregnancy and, hence, the discovery of predictive biomarkers at the pre-diagnostic stages could be of great usefulness in prenatal health management.

8.3.1. Phospholipid profiling of pre-diagnostic preterm delivery and preeclampsia by HILIC-LC/MS

Typical positive-ion HILIC-LC/MS records of pooled 2nd T control plasma showed peaks characteristic of different PL classes at average RT 9.5–10.8 min (PE), 29.5–32.4 min (PC), 34.0–35.9 min (SM) and 38.5–39.8 min (LPC), as shown previously in

Figure 6.6 (Chapter 6). When visually comparing the corresponding MS records obtained for control and disease pools (see Figure 8.6, for PC species), only small changes were suggested for the pre-diagnostic stages of preterm and preeclampsia (arrows in Figure 1a, middle and bottom). A more objective analysis of such complex datasets requires the use of multivariate analysis, used here to probe for changes to be further confirmed rather than for classification (due to the low sample numbers).

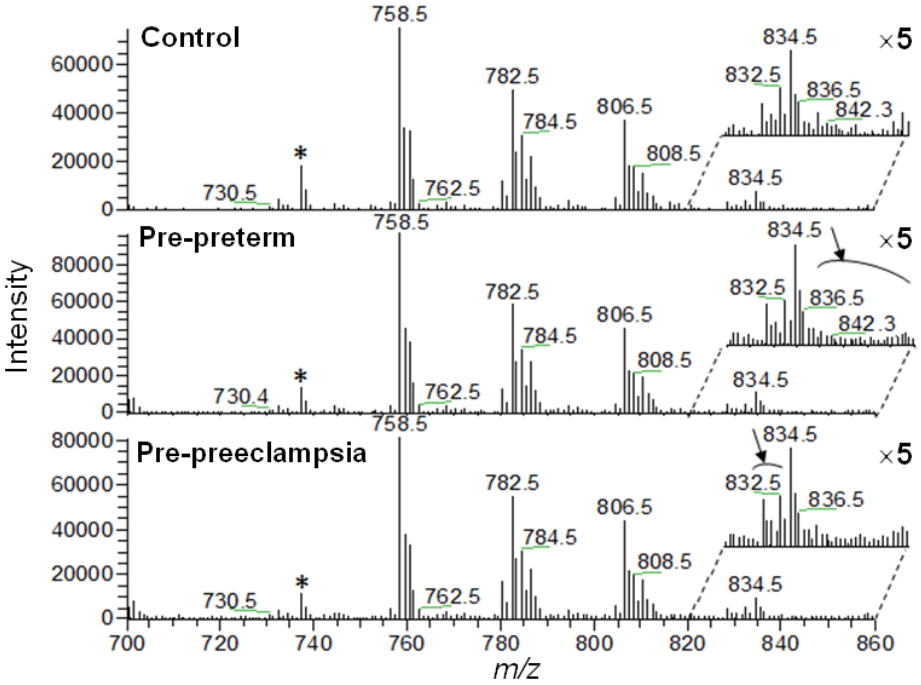


Figure 8.6. a) MS record of PC TIC region (average RT 29.5 – 32.4 min) of pooled plasma samples of pregnant women: healthy throughout pregnancy (control), pre-PTD and pre-PEC. * solvent impurity.

Preliminary PLS-DA (not shown) indicated that each disorder separated from each other (reflecting PL profile specificity) and from controls. Pair-wise PLS-DA (Figure 8.7) produced preliminary models for pre-diagnostic preterm (Q^2 0.70, Figure 8.7a) and preeclampsia (Q^2 0.82, Figure 8.7b). For these disorders, the challenge to find PL markers is raised due to the absence of clinical symptoms, at collection. Loadings (Figure 8.7a and b) inspection revealed lesser numbers of changing features compared to T21 (in Chapter 6, section 6.2). In pre-diagnostic preterm plasma, out of a total of 23 features with VIP $\geq$ 1 and $p < 0.05$, 4 PCs and the *m/z* 842.3 unknown were significantly decreased, along with an increased oxidation product: 1-palmitoyl-2-(9-oxo-nonanoyl)-sn-glycero-3-phosphocholine (PONPC) (Table 8.3, Figure 8.7c left). For pre-diagnostic preeclampsia, decreases were observed for PE(O-36:4) and 4 LPCs (in

common with T21 samples), along with increases in PC(O-40:2) and/or PC(40:9) (Table 8.3, Figure 8.7c right).

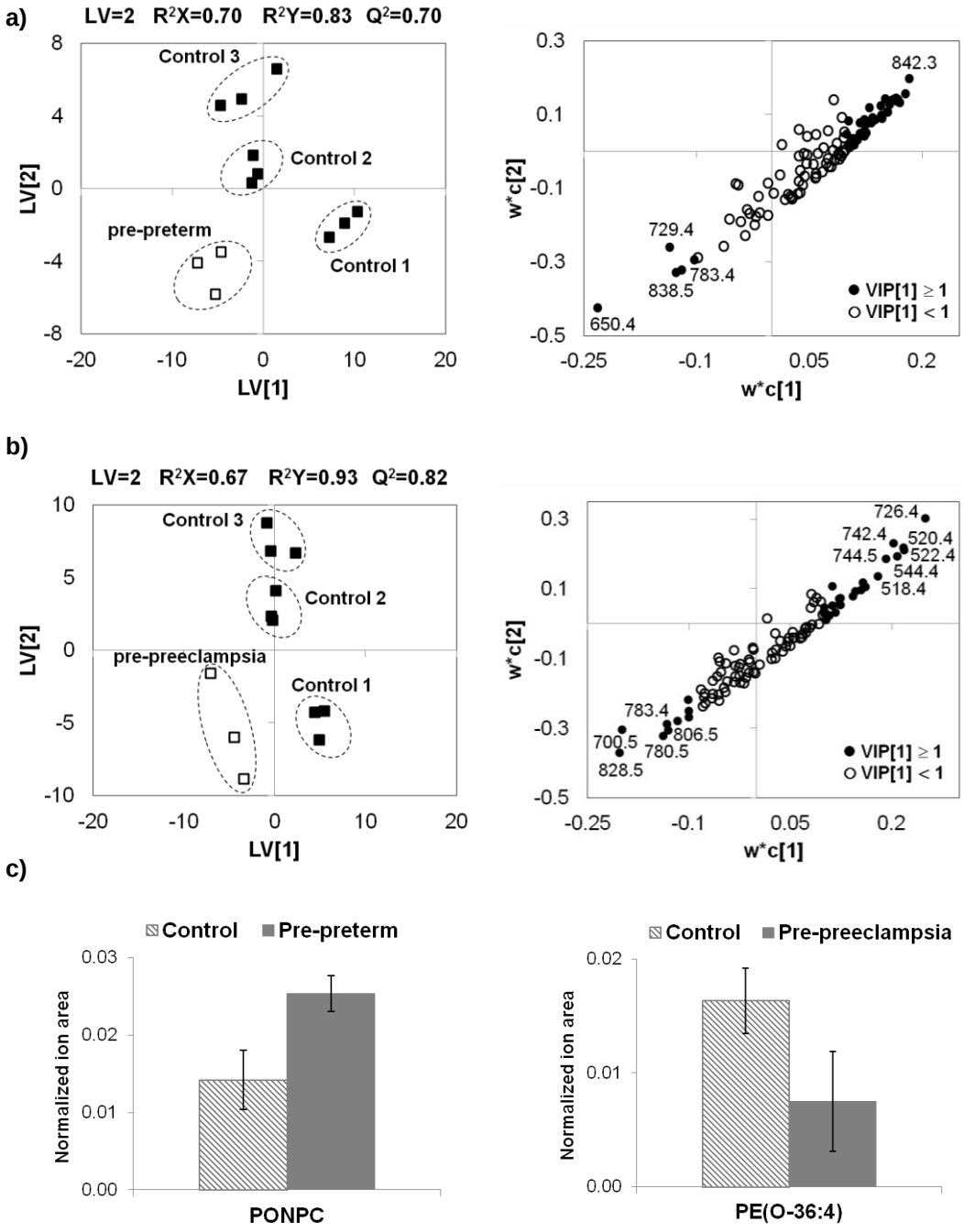


Figure 8.7. PLS-DA scores and loadings plot of pooled plasma samples of a) controls *vs.* pre-diagnostic preterm and b) controls *vs.* pre-diagnostic preeclampsia; c) bar plots of normalized ion area of selected PL for controls vs. pre-diagnostic preterm and vs. pre-diagnostic preeclampsia. PONPC: 1-Palmitoyl-2-(9-oxo-nonanoyl)-sn-glycero-3-phosphocholine (short chain oxidation product of PC(16:0/18:2).

Table 8.3. List of PL species varying most significantly between pre-PTD *vs.* control and pre-PEC *vs.* control samples. Abbreviations: *m/z*: mass-to-charge ratio; RT: retention time; VIP: variable importance to projection, PONPC: 1-Palmitoyl-2-(9-oxo-nonanoyl)-sn-glycero-3-phosphocholine (short chain oxidation product of PC(16:0/18:2)). *unknown compound, † signal observed in more than one disorder (unspecific).

Disease	*m/z*	RT (min)	Identification	VIP[1]	p-value (<0.05)	Relative variation (%)	Effect size
Pre-preterm	650.4	34.4	PONPC	2.35	2.9×10^{-3}	78.5 ± 10.8	2.84 ± 0.84
	842.3	30.9	Unknown * †	1.85	2.5×10^{-2}	-80.3 ± 27.8	-1.72 ± 0.70
	756.5	31.2	PC(34:3)	1.80	2.0×10^{-3}	-24.5 ± 5.8	-1.61 ± 0.69
	784.5	30.8	PC(36:3)	1.71	3.2×10^{-3}	-20.1 ± 5.0	-1.50 ± 0.68
	754.5	31.6	PC(34:4)	1.69	3.8×10^{-3}	-54.5 ± 15.5	-1.47 ± 0.68
	730.5	31.7	PC(32:2)	1.68	6.4×10^{-3}	-37.0 ± 10.7	-1.46 ± 0.68
Pre-preeclampsia	726.4	9.8	PE(O-36:4)	2.52	1.82×10^{-2}	-54.0 ± 19.2	-2.50 ± 0.80
	522.4	39.1	LPC(18:1) †	2.20	3.26×10^{-3}	-37.3 ± 9.1	-1.81 ± 0.71
	520.4	39.4	LPC(18:2) †	2.19	6.77×10^{-3}	-32.8 ± 8.4	-1.80 ± 0.71
	544.4	38.5	LPC(20:4) †	2.09	1.45×10^{-2}	-40.7 ± 11.9	-1.67 ± 0.70
	828.5	30.0	PC(O-40:2) (and PC(40:9))	2.05	8.47×10^{-3}	32.0 ± 7.6	1.61 ± 0.69
	542.3	39.3	LPC(20:5) †	2.04	7.69×10^{-3}	-34.9 ± 9.8	-1.60 ± 0.69

8.3.2. Proposed metabolic interpretation of phospholipid changes

Our results show, for the first time, that the impact on maternal PL is much diminished for the pre-diagnostic stages of preterm and preeclampsia. For pre-diagnostic preterm, unsaturated PC seem particularly affected (decreased), probably due to their high susceptibility to oxidation. The concomitant increase in PONPC, an oxidation product of PC(16:0/18:2), the most abundant PC (*m/z* 758.5 in Figure 8.6), further supports the hypothesis that enhanced oxidation (most probably *via* a radical mechanism (Levitan et al., 2010)) seems also related to a PTD outcome. However, the absence of LPC variations suggests that radical non-enzymatic oxidation may take place, rather than an enzymatic mechanism. Finally, PL sensitivity to pre-diagnostic preeclampsia is confirmed by a number of LPC changes common to T21 (Chapter 6, section 6.2) (possibly reflecting a common oxidative mechanism, a condition suggested before for preeclampsia (Genc et al., 2011)), with specific changes in PE(O-36:4) and an ether-linked PC. Again, although LPC with saturated fatty acyl chains were detected within the plasma lipidome (Figure 8.8), the LPC species that showed significant variations between controls and pre-preeclampsia were those bearing unsaturated PC. LPC comprising unsaturated fatty acyl chains may probably arise from the action of phospholipase A1 or to oxidative cleavage by metal-catalyzed radical oxidation.

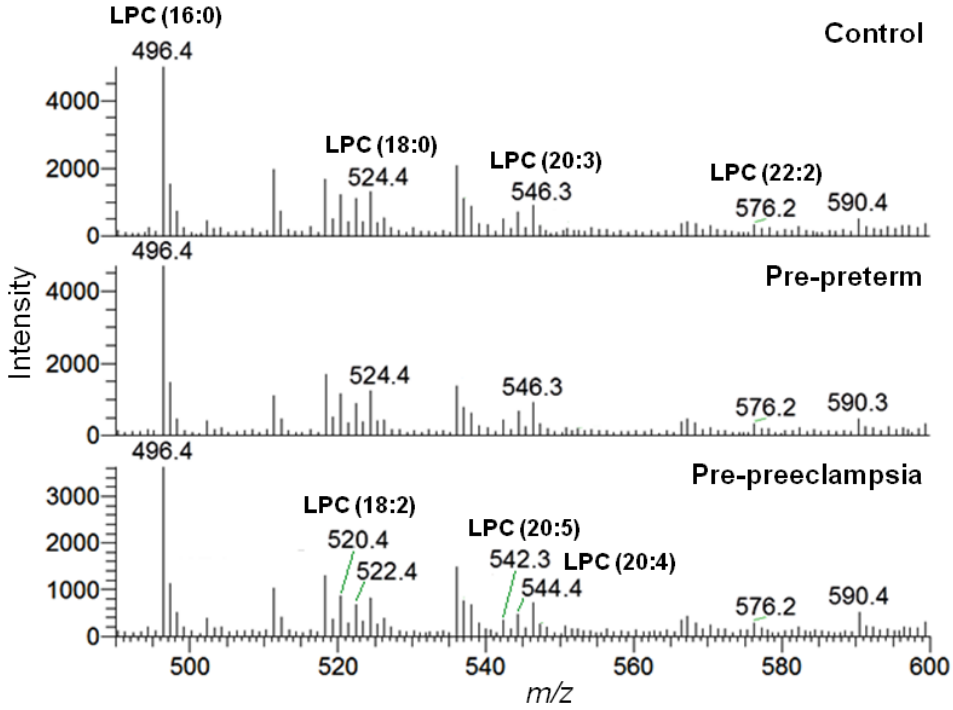


Figure 8.8. MS record of LPC TIC region (average RT 38.5–39.8 min) of pooled plasma samples of pregnant women: healthy throughout pregnancy (control), pre-PTD and pre-PEC.

These results indicated that maternal PL metabolism is exquisitely sensitive to the pre-diagnostic stages of preterm and preeclampsia, although consideration of larger cohorts is necessary for adequate statistical validation to be carried out. The PL marker profiles observed seem specific enough to enable future uses as predictive diagnosis of preterm and preeclampsia to be envisaged, for improved prenatal health management.

Chapter 9. General conclusions and future perspectives

The work presented in this thesis represents a contribution to the study of 1st and 2nd T maternal plasma metabolome and lipidome in order to obtain insight into the biochemistry and pathophysiological alterations occurring in healthy pregnancies and prenatal disorders, mainly using NMR spectroscopy but also MS in selected instances. Since the fundamental role of blood is to maintain body homeostasis, even the smaller alterations found in this work may be of central importance to clinical biochemistry.

Firstly, the human plasma metabolome was characterized by 1D and 2D NMR spectroscopy enabling the identification of 31 low M_w metabolites (including 1,5-anhydroglucitol, here detected for the first time to our knowledge) in addition to lipids and proteins (albumin and N-acetyl glycoproteins). The characterization of plasma lipidome by NMR spectroscopy has been less explored in literature, and here the assignment of plasma lipid extracts by 1D and 2D NMR enabled the identification of four distinct unsaturated fatty acids, total saturated fatty acids, free cholesterol, esterified cholesterol, total phospholipids, phosphatidylcholines, sphingomyelins and triglycerides, confirming previous reports. Furthermore, STOCSY analysis gave additional information mainly in the identification of different CH_3 environments arising from distinct unsaturated FA. MQ NMR spectroscopy was explored in this work for assignment of complex lipid mixtures for the first time to our knowledge. This technique showed simpler spectral profiles than single quantum 2D NMR experiments, however the molecules elected to test this methodology showed a similar fingerprint with high overlap of protons from fatty acyl chain and from cholesterol part of EC in the aliphatic region. The application of MQ to plasma lipid extracts unveiled a similar profile when compared to the lipid mixture. In future, the combination of MQ with diffusion ordered spectroscopy (DOSY) may allow the discrimination of similar lipid molecules by separation of the MQ fingerprint as a consequence of molecular size.

The identification of the effects of handling and storage conditions on plasma NMR metabolome was of paramount importance for the proposed metabolic profiling studies of healthy pregnancy and prenatal disorders. EDTA and heparin tubes seem equally suitable for plasma collection for NMR analysis, however EDTA signals overlap with some resonances that are not easily studied in other spectral regions (choline containing compounds PC, LPC and SM). Room temperature stability studies showed that plasma samples should not be kept at room temperature for $\geq$ 2.5 hours, since metabolic changes alter the plasma composition, particularly in relation to lipoprotein profile. Short term (up to 1 month) storage stability revealed that plasma samples can be stored at -20ºC up to 7 days, but for longer periods they should be stored at -80ºC. Long term storage at -80ºC showed almost negligible alterations up to 30 months with only a small increase in cholesterol having been noted after 20 months storage. The

effects of up to 5 freeze-thaw cycles were found to be strongly sample-dependent with most alterations seen to take place upon cycle 4 and thereafter, hence no more than 3 freeze-thaw cycles are recommended. Finally, the particular non-fasting conditions characterizing samples collection for the pregnant women in this work (2 hours after ingestion of uncontrolled breakfast) were evaluated in a group of healthy non-pregnant women, showing that non-fasting conditions do not affect the overall metabolic characteristics of the sample group.

The second study aimed at defining and characterizing the metabolic adaptations reflected in plasma metabolome throughout healthy pregnancies, constituting an important step towards the identification of disease-related deviations in subsequent studies. Results showed the expected decrease in circulating amino acids early in pregnancy, suggesting a larger requirement by the fetus. Newly observed changes included early adjustments in energy and gut microflora metabolisms. In addition, alterations in glomerular filtration rate were observed, along with HDL and LDL+VLDL increases throughout pregnancy in different extents and accompanied by increases in fatty acid chain length and degree of unsaturation. Plasma and urine metabolite correlation studies provided further understanding of healthy pregnancy metabolism, namely unveiling specific lipoprotein/protein plasma metabolic correlations with excreted metabolites.

NMR metabolomics of 1st and 2nd trimester maternal plasma of a general CD group and a T21 subgroup, compared to controls, showed additional information accompanying these disorders. A different plasma metabolic profile of CD and T21 was observed in the 2nd trimester compared to 1st, suggesting a dynamic metabolic evolution in CD cases. The metabolic pattern of enhanced β-oxidation observed previously in the 1st trimester was maintained in the 2nd trimester, novel changes including abnormalities in plasma lipoprotein profile and its involvement in β-oxidation, along with possible changes in fetal renal function. Additionally, plasma lipid extracts gave valuable insight into the mediation of anti-inflammatory and pro-inflammatory processes indicating a more significant response to inflammation for 2nd trimester T21 cases. Plasma/urine correlation studies confirmed main changes previously noted for each biofluid, unveiling possible relationships relating to plasma creatine and lipids (particularly LDL+VLDL); excreted ketone bodies and acetate/formate; allantoin and sugar metabolism; and gut microflora metabolites. Furthermore, HILIC-LC/MS analysis of phospholipid extracts of 2nd trimester T21 cases showed alterations in levels of LPC, ether-linked PC and PE consistent with a possible oxidative stress accompanying T21. These results demonstrate the potential of plasma metabolomics in monitoring and characterizing CD and T21 cases, however, validation in larger cohorts and

demonstration of putative biochemical hypothesis are required before clinical applications may be envisaged. In addition, the metabolic signatures of T18 and T13 require investigation in order to define a specific metabolic signature capable to discriminate each trisomy type.

The possible application of NMR plasma metabolomics for GDM management was investigated considering 2nd trimester maternal plasma and corresponding lipid extracts obtained for a group of pregnant women without clinical signs but who later developed GDM (pre-diagnosis group) and a group of pregnant women at the time of GDM diagnosis (post-diagnosis group), in comparison to controls. Plasma and lipid extracts displayed comparable classification performance, the former enabling direct and more rapid analysis. Metabolic variations detectable in pre-diagnosis GDM included an increase in free and/or esterified cholesterol in plasma, accompanied by smaller increases in plasma HDL, LDL+VLDL lipoproteins, and changes in FA (e.g., 18:2 and 20:4) and TG levels viewed through lipid extracts. Novel alterations were observed for glucose (small increase), some glycolytic and TCA cycle metabolites, along with a possible deregulation of gut microflora, homocysteine-methionine conversion and urea cycle. Many of the changes observed in the pre-diagnosis GDM were intensified at the time of diagnosis, whereas new variations were noted for some small metabolites. In addition, for post-diagnosis GDM, PLS regression identified metabolites with higher dependence of normal gestational age evolution whose were also compared with the metabolic trajectory previously defined for healthy pregnancies. After exclusion of plasma metabolic changes with a possible contribution of gestational age, a robust profile of 26 resonances may be used to discriminate GDM. This work advances the possibility of predicting GDM with basis on the full resolution NMR spectra of plasma and using a multi-metabolite biomarker of GDM as a complementary tool in the clinical management of the disease. In future studies, the putative metabolic disturbances hereby suggested to relate to pre- and post-diagnosis GDM may be demonstrated in larger cohorts of distinct ethnicities and by complementary biochemical methods.

Prenatal disorders with less impact on 2nd T maternal plasma metabolome included FM and pre-diagnosis PROM, preterm delivery and preeclampsia. Metabolic alterations found for maternal plasma of pregnant women carrying malformed fetuses were consistent with a previous report with some novel alterations found, namely LDL+VLDL lipid composition, some low M_w metabolites and albumin. Overall these changes suggested increased β-oxidation of fatty acids, depletion of phospholipids in LDL+VLDL particles, enhanced glycolysis and protein biosynthesis and possible alterations in urea cycle and gut microflora. CNS malformations showed a distinct metabolic profile when compared to other FM types suggesting enhanced glycerolipid

metabolism, decreased N-acetyl glycoproteins metabolism and a possible lower contribution of CNS for the β-oxidation, enhanced glycolysis and urea cycle alterations found for general FM cases.

The possibility of predicting the occurrence of PROM, occurring after 37th g.w., was attempted 19-24 g.w. before PROM diagnosis by NMR metabolomics. This unveiled a specific metabolic signature characterized by an increased biosynthesis of N-acetyl glycoproteins, along with subtle alterations in lipid profile, amino acid catabolism and urea cycle. Despite some detected alterations, any impact of this disorder on plasma profile is surprising due to the large time (19-24 g.w.) between sample collection and PROM event. Thus, the results obtained really need confirmation in larger cohorts.

NMR metabolomics of pre-diagnostic preterm delivery unveiled no impact of this disorder on plasma metabolome detectable by NMR, while for preeclampsia, an insuffient number of samples was available. The HILIC-LC/MS phospholipid profile of 2nd T maternal plasma showed apparent specific alterations for prediction of preterm delivery and preeclampsia. Pre-diagnostic preterm delivery showed alterations in unsaturated PC and a PC oxidation product, possible due to non-enzymatic oxidative processes, while phospholipid sensitivity was confirmed for pre-diagnostic preeclampsia by changes in LPC (common to T21, possibly reflecting a common oxidative mechanism) and novel changes in PE and an ether-linked PC. These results indicated that 2nd T maternal plasma is sensitive to FM and pre-diagnosis PROM by NMR, and to preterm delivery and preeclampsia by HILIC-LC/MS profiling of phospholipids. The apparent specificity of metabolic profiles requires further investigation in larger cohorts before future use for complementary or predictive diagnosis may be advanced.

The results presented in this thesis give new insights into the development of new prenatal diagnosis and prognosis methods through maternal plasma metabolome. Future work will entail the study of other important prenatal disorders, not studied in this work due the low number of samples available, namely IUGR, SGA and macrosomia. Furthermore, larger cohorts and samples from different ethnicities will be required to confirm the changes found in this work, especially for the most serious disorders namely trissomies, and prediction of GDM, preeclampsia and preterm delivery. Another future work of paramount importance includes external validation by classification of new samples using the previously developed MVA models in order to advance a possible clinical application of maternal plasma biomarkers. Finally, the search for potential biomarkers for diagnosis and prediction of prenatal disorders will

necessarily entail blood samples collected early, in the 1st trimester, in order to more significantly reduce the adverse risks caused to mother and fetus.

References

Ala-korpela, M.; Korhonen, A.; Keisala, J.; Horkko, S.; Korpi, P.; Ingman, L. P.; Jokisaari, J.; Savolainen, M. J.; Kesaniemi, Y. A. ^{1}H NMR-Based Absolute Quantitation of Human Lipoproteins and Their Lipid Contents Directly from Plasma. *Journal of Lipid Research* **1994**, *35*, 2292–2304.

Ala-Korpela, M.; Lankinen, N.; Salminen, A.; Suna, T.; Soininen, P.; Laatikainen, R.; Ingman, P.; Jauhiainen, M.; Taskinen, M. R.; Heberger, K.; et al. The Inherent Accuracy of ^{1}H NMR Spectroscopy to Quantify Plasma Lipoproteins Is Subclass Dependent. *Atherosclerosis* **2007**, *190*, 352–358.

Alexandre-Gouabau, M.-C.; Courant, F.; Moyon, T.; Küster, A.; Le Gall, G.; Tea, I.; Antignac, J.-P.; Darmaun, D. Maternal and Cord Blood LC-HRMS Metabolomics Reveal Alterations in Energy and Polyamine Metabolism, and Oxidative Stress in Very-Low Birth Weight Infants. *Journal of Proteome Research* **2013**, *12*, 2764–2778.

Almeida, L. A. M. Lipidómica do Plasma Materno para o Diagnóstico de Diabetes Gestacional e Trissomia 21, University of Aveiro, 2014.

Amorini, A. M.; Giorlandino, C.; Longo, S.; D'Urso, S.; Mesoraca, A.; Santoro, M. L.; Picardi, M.; Gullotta, S.; Cignini, P.; Lazzarino, D.; et al. Metabolic Profile of Amniotic Fluid as a Biochemical Tool to Screen for Inborn Errors of Metabolism and Fetal Anomalies. *Molecular and Cellular Biochemistry* **2012**, *359*, 205–216.

Andersen, C. M.; Bro, R. Variable Selection in Regression-a Tutorial. *Journal of Chemometrics* **2010**, *24*, 728–737.

Atzori, L.; Antonucci, R.; Barberini, L.; Locci, E.; Marincola, F. C.; Scano, P.; Cortesi, P.; Agostiniani, R.; Defraia, R.; Weljie, A.; et al. ^{1}H NMR-Based Metabolomic Analysis of Urine from Preterm and Term Neonates. *Frontiers in Bioscience (Elite edition)* **2011**, *3*, 1005–1012.

Auray-Blais, C.; Raiche, E.; Gagnon, R.; Berthiaume, M.; Pasquier, J.-C. Metabolomics and Preterm Birth: What Biomarkers in Cervicovaginal Secretions Are Predictive of High-Risk Pregnant Women? *International Journal of Mass Spectrometry* **2011**, *307*, 33–38.

Austdal, M.; Skråstad, R. B.; Gundersen, A. S.; Austgulen, R.; Iversen, A.-C.; Bathen, T. F. Metabolomic Biomarkers in Serum and Urine in Women with Preeclampsia. *PLoS ONE* **2014**, *9*, e91923.

Bacq, Y.; Zarka, O.; Bréchot, J. F.; Mariotte, N.; Vol, S.; Tichet, J.; Weill, J. Liver Function Tests in Normal Pregnancy: A Prospective Study of 103 Pregnant Women and 103 Matched Controls. *Hepatology (Baltimore, Md.)* **1996**, *23*, 1030–1034.

Bahado-Singh, R. O.; Akolekar, R.; Mandal, R.; Dong, E.; Xia, J.; Kruger, M.; Wishart, D. S.; Nicolaides, K. Metabolomics and First-Trimester Prediction of Early-Onset Preeclampsia. *The Journal of Maternal-Fetal & Neonatal Medicine* **2012**, *25*, 1840–1847.

Bahado-Singh, R. O.; Akolekar, R.; Mandal, R.; Dong, E.; Xia, J.; Kruger, M.; Wishart, D. S.; Nicolaides, K. First-Trimester Metabolomic Detection of Late-Onset Preeclampsia. *American Journal of Obstetrics and Gynecology* **2013a**, *208*, 58.e1–e7.

Bahado-Singh, R. O.; Akolekar, R.; Mandal, R.; Dong, E.; Xia, J.; Kruger, M.; Wishart, D. S.; Nicolaides, K. Metabolomic Analysis for First-Trimester Down Syndrome Prediction. *American Journal of Obstetrics and Gynecology* **2013b**, *208*, 371.e1–e8.

Bahado-Singh, R. O.; Akolekar, R.; Chelliah, A.; Mandal, R.; Dong, E.; Kruger, M.; Wishart, D. S.; Nicolaides, K. Metabolomic Analysis for First-Trimester Trisomy 18 Detection. *American Journal of Obstetrics and Gynecology* **2013c**, *209*, 65.e1–e9.

Bahado-Singh, R. O.; Ertl, R.; Mandal, R.; Bjorndahl, T. C.; Syngelaki, A.; Han, B.; Dong, E.; Liu, P. B.; Alpay-Savasan, Z.; Wishart, D. S.; et al. Metabolomic Prediction of Fetal Congenital Heart Defect in the First Trimester. *American Journal of Obstetrics and Gynecology* **2014**, *211*, 240.e1–e240.e14.

Baig, S.; Lim, J. Y.; Fernandis, A. Z.; Wenk, M. R.; Kale, A.; Su, L. L.; Biswas, A.; Vasoo, S.; Shui, G.; Choolani, M. Lipidomic Analysis of Human Placental Syncytiotrophoblast Microvesicles in Adverse Pregnancy Outcomes. *Placenta* **2013**, *34*, 436–442.

Balci, M. *Basic ^{1}H- and ^{13}C-NMR Spectroscopy*; Elsevier, 2005.

Bamfo, J. E. A. K.; Odibo, A. O. Diagnosis and Management of Fetal Growth Restriction. *Journal of Pregnancy* **2011**, *2011*.

Barker, M.; Rayens, W. Partial Least Squares for Discrimination. *Journal of Chemometrics* **2003**, *17*, 166–173.

Barri, T.; Dragsted, L. O. UPLC-ESI-QTOF/MS and Multivariate Data Analysis for Blood Plasma and Serum Metabolomics: Effect of Experimental Artefacts and Anticoagulant. *Analytica Chimica Acta* **2013**, *768*, 118–128.

Barton, R. H.; Nicholson, J. K.; Elliott, P.; Holmes, E. High-Throughput ^{1}H NMR-Based Metabolic Analysis of Human Serum and Urine for Large-Scale Epidemiological Studies: Validation Study. *International Journal of Epidemiology* **2008**, *37*, i31–i40.

Barton, R. H.; Waterman, D.; Bonner, F. W.; Holmes, E.; Clarke, R.; the, P. C.; Nicholson, J. K.; Lindon, J. C. The Influence of EDTA and Citrate Anticoagulant Addition to Human Plasma on Information Recovery from NMR-Based Metabolic Profiling Studies. *Molecular BioSystems* **2010**, *6*, 215–224.

Bastian, M.; Heymann, S.; Jacomy, M. Gephi: An Open Source Software for Exploring and Manipulating Network. In; International AAAI Conference on Weblogs and Social Media, 2009.

Bax, A.; Davis, D. G. MLEV-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy. *Journal of Magnetic Resonance (1969)* **1985**, *65*, 355–360.

Beck, S.; Wojdyla, D.; Say, L.; Betran, A. P.; Merialdi, M.; Requejo, J. H.; Rubens, C.; Menon, R.; Van Look, P. F. A. The Worldwide Incidence of Preterm Birth: A Systematic Review of Maternal Mortality and Morbidity. *Bulletin of the World Health Organization* **2010**, *88*, 31–38.

Beckonert, O.; Keun, H. C.; Ebbels, T. M. D.; Bundy, J.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. Metabolic Profiling, Metabolomic and Metabonomic Procedures for NMR Spectroscopy of Urine, Plasma, Serum and Tissue Extracts. *Nature Protocols* **2007**, *2*, 2692–2703.

Benn, P. A.; Kaminsky, L. M.; Ying, J.; Borgida, A. F.; Egan, J. F. X. Combined Second-Trimester Biochemical and Ultrasound Screening for Down Syndrome. *Obstetrics and Gynecology* **2002**, *100*, 1168–1176.

Benn, P.; Cuckle, H.; Pergament, E. Non-Invasive Prenatal Testing for Aneuploidy: Current Status and Future Prospects. *Ultrasound in obstetrics & gynecology* **2013**, *42*, 15–33.

Bentley-Lewis, R.; Huynh, J.; Xiong, G.; Lee, H.; Wenger, J.; Clish, C.; Nathan, D.; Thadhani, R.; Gerszten, R. Metabolomic Profiling in the Prediction of Gestational Diabetes Mellitus. *Diabetologia* **2015**, DOI: 10.1007/s00125-015-3553-4.

Berben, L.; Sereika, S. M.; Engberg, S. Effect Size Estimation: Methods and Examples. *International Journal of Nursing Studies* **2012**, *49*, 1039–1047.

Bernini, P.; Bertini, I.; Luchinat, C.; Nincheri, P.; Staderini, S.; Turano, P. Standard Operating Procedures for Pre-Analytical Handling of Blood and Urine for Metabolomic Studies and Biobanks. *Journal of Biomolecular NMR* **2011**, *49*, 231–243.

Blekherman, G.; Laubenbacher, R.; Cortes, D. F.; Mendes, P.; Torti, F. M.; Akman, S.; Torti, S. V; Shulaev, V. Bioinformatics Tools for Cancer Metabolomics. *Metabolomics* **2011**, *7*, 329–343.

Blencowe, H.; Cousens, S.; Chou, D.; Oestergaard, M.; Say, L.; Moller, A.-B.; Kinney, M.; Lawn, J. Born Too Soon: The Global Epidemiology of 15 Million Preterm Births. *Reproductive Health* **2013**, *10 Suppl 1*, S2.

Bock, J. L. Metabolic Profiling of Amniotic Fluid by Proton Nuclear Magnetic Resonance Spectroscopy: Correlation with Fetal Maturation and Other Clinical Variables. *Clinical Chemistry* **1994**, *40*, 56–61.

Bonvallot, N.; Tremblay-Franco, M.; Chevrier, C.; Canlet, C.; Warembourg, C.; Cravedi, J.-P.; Cordier, S. Metabolomics Tools for Describing Complex Pesticide Exposure in Pregnant Women in Brittany (France). *PLoS ONE* **2013**, *8*, e64433.

Bredaki, F. E.; Poon, L. C.; Birdir, C.; Escalante, D.; Nicolaides, K. H. First-Trimester Screening for Neural Tube Defects Using Alpha-Fetoprotein. *Fetal Diagnosis and Therapy* **2012**, *31*, 109–114.

Broadhurst, D.; Kell, D. Statistical Strategies for Avoiding False Discoveries in Metabolomics and Related Experiments. *Metabolomics* **2006**, *2*, 171–196.

Butte, N. F. Carbohydrate and Lipid Metabolism in Pregnancy: Normal Compared with Gestational Diabetes Mellitus. *American Journal of Clinical Nutrition* **2000**, *71*, 1256S – 1261S.

Cetin, I.; Alvino, G. Intrauterine Growth Restriction: Implications for Placental Metabolism and Transport. A Review. *Placenta* **2009**, *30 Suppl A*, S77–S82.

Chen, X. H.; Scholl, T. O.; Leskiw, M.; Savaille, J.; Stein, T. P. Differences in Maternal Circulating Fatty Acid Composition and Dietary Fat Intake in Women With Gestational Diabetes Mellitus or Mild Gestational Hyperglycemia. *Diabetes Care* **2010**, *33*, 2049–2054.

Cheung, K. L.; Lafayette, R. A. Renal Physiology of Pregnancy. *Advances in Chronic Kidney Disease* **2013**, *20*, 209–214.

Chi, Y.; Pei, L.; Chen, G.; Song, X.; Zhao, A.; Chen, T.; Su, M.; Zhang, Y.; Liu, J.; Ren, A.; et al. Metabonomic Profiling of Human Placentas Reveals Different Metabolic Patterns among Subtypes of Neural Tube Defects. *Journal of Proteome Research* **2014**, *13*, 934–945.

Christensen, B. L.; Williams, M. Assessing Postprandial Glucose Using 1,5-Anhydroglucitol: An Integrative Literature Review. *Journal of the American Academy of Nurse Practitioners* **2009**, *21*, 542–548.

Christie, W. W. Sterols 1. Cholesterol and cholesterol esters: structure, occurrence, biochemistry and analysis http://lipidlibrary.aocs.org/Lipids/cholest/index.htm (accessed Apr 1, 2015).

Ciborowski, M.; Zbucka-Kretowska, M.; Bomba-Opon, D.; Wielgos, M.; Brawura-Biskupski-Samaha, R.; Pierzynski, P.; Szmitkowski, M.; Wolczynski, S.; Lipinska, D.; Citko, A.; et al. Potential First Trimester Metabolomic Biomarkers of Abnormal Birth Weight in Healthy Pregnancies. *Prenatal Diagnosis* **2014**, *34*, 870–877.

Cífková, E.; Holčapek, M.; Lísa, M.; Ovčačíková, M.; Lyčka, A.; Lynen, F.; Sandra, P. Nontargeted Quantitation of Lipid Classes Using Hydrophilic Interaction Liquid Chromatography-Electrospray Ionization Mass Spectrometry with Single Internal Standard and Response Factor Approach. *Analytical Chemistry* **2012**, *84*, 10064–10070.

Claridge, T. D. W. *High-Resolution NMR Techniques in Organic Chemistry*, 2nd ed.; Elsevier: Oxford, 2009.

Cloarec, O.; Dumas, M.-E.; Craig, A.; Barton, R. H.; Trygg, J.; Hudson, J.; Blancher, C.; Gauguier, D.; Lindon, J. C.; Holmes, E.; et al. Statistical Total Correlation Spectroscopy: An Exploratory Approach for Latent Biomarker Identification from Metabolic ^{1}H NMR Data Sets. *Analytical Chemistry* **2005**, *77*, 1282–1289.

Coen, M.; Hong, Y.-S.; Cloarec, O.; Rhode, C. M.; Reily, M. D.; Robertson, D. G.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. Heteronuclear ^{1}H-^{31}P Statistical Total Correlation NMR Spectroscopy of Intact Liver for Metabolic Biomarker Assignment: Application to Galactosamine-Induced Hepatotoxicity. *Analytical Chemistry* **2007**, *79*, 8956–8966.

Coen, M.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. NMR-Based Metabolic Profiling and Metabonomic Approaches to Problems in Molecular Toxicology. *Chemical Research in Toxicology* **2008**, *21*, 9–27.

Cohen, Y.; Avram, L.; Frish, L. Diffusion NMR Spectroscopy in Supramolecular and Combinatorial Chemistry: An Old Parameter-New Insights. *Angewandte Chemie (International ed. in English)* **2005**, *44*, 520–554.

Cohn, B. R.; Joe, B. N.; Zhao, S. J.; Kornak, J.; Zhang, V. Y.; Iman, R.; Kurhanewicz, J.; Vahidi, K.; Yu, J. W.; Caughey, A. B.; et al. Quantitative Metabolic Profiles of 2nd and 3rd Trimester Human Amniotic Fluid Using ^{1}H HR-MAS Spectroscopy. *Magnetic Resonance Materials in Physics Biology and Medicine* **2009**, *22*, 343–352.

Cohn, B. R.; Fukuchi, E. Y.; Joe, B. N.; Swanson, M. G.; Kurhanewicz, J.; Yu, J. W.; Caughey, A. B. Calculation of Gestational Age in Late Second and Third Trimesters by Ex Vivo Magnetic Resonance Spectroscopy of Amniotic Fluid. *American Journal of Obstetrics and Gynecology* **2010**, *203*, 76.e1–e76.e10.

Collino, S.; Martin, F.-P.; Kochhar, S.; Rezzi, S. Monitoring Healthy Metabolic Trajectories with Nutritional Metabonomics. *Nutrients* **2009**, *1*, 101–110.

Collins, S. L.; Impey, L. Prenatal Diagnosis: Types and Techniques. *Early Human Development* **2012**, *88*, 3–8.

Cosmi, E.; Visentin, S.; Favretto, D.; Tucci, M.; Ragazzi, E.; Viel, G.; Ferrara, S. D. Selective Intrauterine Growth Restriction in Monochorionic Twin Pregnancies: Markers of Endothelial Damage and Metabolomic Profile. *Twin Research and Human Genetics Studies* **2013**, *16*, 816–826.

Creasy, R. K.; Resnik, R.; Iams, J. D. *Creasy and Resnik's Maternal-Fetal Medicine: Principles and Practice*, 6th ed.; Saunders Elsevier: Philadelphia, 2009.

Cui, X.; Churchill, G. A. Statistical Tests for Differential Expression in cDNA Microarray Experiments. *Genome Biology* **2003**, *4*, 210.

Dallüge, J.; Beens, J.; Brinkman, U. A. T. Comprehensive Two-Dimensional Gas Chromatography: A Powerful and Versatile Analytical Tool. *Journal of Chromatography A* **2003**, *1000*, 69–108.

Dani, C.; Bresci, C.; Berti, E.; Ottanelli, S.; Mello, G.; Mecacci, F.; Breschi, R.; Hu, X.; Tenori, L.; Luchinat, C. Metabolomic Profile of Term Infants of Gestational Diabetic Mothers. *The Journal of Maternal-Fetal & Neonatal Medicine* **2014**, *27*, 537–542.

Dasarathy, J.; Gruca, L. L.; Bennett, C.; Parimi, P. S.; Duenas, C.; Marczewski, S.; Fierro, J. L.; Kalhan, S. C. Methionine Metabolism in Human Pregnancy. *American Journal of Clinical Nutrition* **2010**, *91*, 357–365.

Dénes, J.; Szabo, E.; Robinette, S. L.; Szatmari, I.; Szonyi, L.; Kreuder, J. G.; Rauterberg, E. W.; Takats, Z. Metabonomics of Newborn Screening Dried Blood Spot Samples: A Novel Approach in the Screening and Diagnostics of Inborn Errors of Metabolism. *Analytical Chemistry* **2012**, *84*, 10113–10120.

Deprez, S.; Sweatman, B. C.; Connor, S. C.; Haselden, J. N.; Waterfield, C. J. Optimisation of Collection, Storage and Preparation of Rat Plasma for ^{1}H NMR Spectroscopic Analysis in Toxicology Studies to Determine Inherent Variation in Biochemical Profiles. *Journal of Pharmaceutical and Biomedical Analysis* **2002**, *30*, 1297–1310.

DeSisto, C. L.; Kim, S. Y.; Sharma, A. J. Prevalence Estimates of Gestational Diabetes Mellitus in the United States, Pregnancy Risk Assessment Monitoring System (PRAMS), 2007-2010. *Preventing Chronic Disease* **2014**, *11*, E104.

Dessì, A.; Atzori, L.; Noto, A.; Visser, G. H. A.; Gazzolo, D.; Zanardo, V.; Barberini, L.; Puddu, M.; Ottonello, G.; Atzei, A.; et al. Metabolomics in Newborns with Intrauterine Growth Retardation (IUGR): Urine Reveals Markers of Metabolic Syndrome. *The Journal of Maternal-Fetal & Neonatal Medicine* **2011**, *24 Suppl 2*, 35–39.

Dessì, A.; Ottonello, G.; Fanos, V.; Dessi, A.; Ottonello, G.; Fanos, V. Physiopathology of Intrauterine Growth Retardation: From Classic Data to Metabolomics. *Journal of Maternal-Fetal & Neonatal Medicine* **2012**, *25*, 13–18.

Dessì, A.; Puddu, M.; Ottonello, G.; Fanos, V. Metabolomics and Fetal-Neonatal Nutrition: Between "Not Enough" and "Too Much". *Molecules* **2013**, *18*, 11724–11732.

Dessì, A.; Marincola, F. C. F. C.; Fanos, V. Metabolomics and the Great Obstetrical Syndromes – GDM, PET and IUGR. *Best Practice & Research Clinical Obstetrics & Gynaecology* **2014a**, 1–9.

Dessì, A.; Corsello, G.; Stronati, M.; Gazzolo, D.; Caboni, P.; Carboni, R.; Fanos, V. New Diagnostic Possibilities in Systemic Neonatal Infections: Metabolomics. *Early Human Development* **2014b**, *90 Suppl 1*, S19–S21.

Diaz, S. O.; Pinto, J.; Graça, G.; Duarte, I. F.; Barros, A. S.; Galhano, E.; Pita, C.; Almeida, M. do C.; Goodfellow, B. J.; Carreira, I. M.; et al. Metabolic Biomarkers of Prenatal Disorders: An Exploratory NMR Metabonomics Study of Second Trimester Maternal Urine and Blood Plasma. *Journal of Proteome Research* **2011**, *10*, 3732–3742.

Diaz, S. O.; Barros, A. S.; Goodfellow, B. J.; Duarte, I. F.; Carreira, I. M.; Galhano, E.; Pita, C.; Almeida, M. do C.; Gil, A. M. Following Healthy Pregnancy by Nuclear Magnetic Resonance (NMR) Metabolic Profiling of Human Urine. *Journal of Proteome Research* **2013a**, *12*, 969–979.

Diaz, S. O.; Barros, A. S.; Goodfellow, B. J.; Duarte, I. F.; Galhano, E.; Pita, C.; Almeida, M. do C.; Carreira, I. M.; Gil, A. M. Second Trimester Maternal Urine for the Diagnosis of Trisomy 21 and Prediction of Poor Pregnancy Outcomes. *Journal of Proteome Research* **2013b**, *12*, 2946–2957.

Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H. Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in ^{1}H NMR Metabonomics. *Analytical Chemistry* **2006**, *78*, 4281–4290.

Duarte, I. F.; Gil, A. M. Metabolic Signatures of Cancer Unveiled by NMR Spectroscopy of Human Biofluids. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2012**, *62*, 51–74.

Duarte, I. F.; Rocha, C. M.; Gil, A. M. Metabolic Profiling of Biofluids: Potential in Lung Cancer Screening and Diagnosis. *Expert Review of Molecular Diagnostics* **2013**, *13*, 737–748.

Dudzik, D.; Zorawski, M.; Skotnicki, M.; Zarzycki, W.; Kozlowska, G.; Bibik-Malinowska, K.; Vallejo, M.; García, A.; Barbas, C.; Ramos, M. P. Metabolic Fingerprint of Gestational Diabetes Mellitus. *Journal of Proteomics* **2014**, *103*, 57–71.

Dumas, M.-E.; Wilder, S. P.; Bihoreau, M.-T.; Barton, R. H.; Fearnside, J. F.; Argoud, K.; D'Amato, L.; Wallis, R. H.; Blancher, C.; Keun, H. C.; et al. Direct Quantitative Trait Locus Mapping of Mammalian Metabolic Phenotypes in Diabetic and Normoglycemic Rat Models. *Nature Genetics* **2007**, *39*, 666–672.

Dunn, W. B.; Broadhurst, D.; Ellis, D. I.; Brown, M.; Halsall, A.; O'Hagan, S.; Spasic, I.; Tseng, A.; Kell, D. B. A GC-TOF-MS Study of the Stability of Serum and Urine Metabolomes during the UK Biobank Sample Collection and Preparation Protocols. *International Journal of Epidemiology* **2008**, *37*, i23–i30.

Dunn, W. B.; Brown, M.; Worton, S. A.; Crocker, I. P.; Broadhurst, D.; Horgan, R.; Kenny, L. C.; Baker, P. N.; Kell, D. B.; Heazell, A. E. P. Changes in the Metabolic Footprint of Placental Explant-Conditioned Culture Medium Identifies Metabolic Disturbances Related to Hypoxia and Pre-Eclampsia. *Placenta* **2009**, *30*, 974–980.

Dunn, W. B.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-McIntyre, S.; Anderson, N.; Brown, M.; Knowles, J. D.; Halsall, A.; Haselden, J. N.; et al. Procedures for Large-Scale Metabolic Profiling of Serum and Plasma Using Gas Chromatography and Liquid Chromatography Coupled to Mass Spectrometry. *Nature Protocols* **2011a**, *6*, 1060–1083.

Dunn, W. B.; Brown, M.; Worton, S. A.; Davies, K.; Jones, R. L.; Kell, D. B.; Heazell, A. E. P. The Metabolome of Human Placental Tissue: Investigation of First Trimester Tissue and Changes Related to Preeclampsia in Late Pregnancy. *Metabolomics* **2011b**, *8*, 579–597.

Edelstam, G.; Karlsson, C.; Westgren, M.; Löwbeer, C.; Swahn, M.-L. Human Chorionic Gonadatropin (hCG) during Third Trimester Pregnancy. *Scandinavian Journal of Clinical and Laboratory Investigation* **2007**, *67*, 519–525.

Ellis, D. I.; Goodacre, R. Metabolic Fingerprinting in Disease Diagnosis: Biomedical Applications of Infrared and Raman Spectroscopy. *The Analyst* **2006**, *131*, 875–885.

Emet, T.; Ustüner, I.; Güven, S. G.; Balık, G.; Ural, U. M.; Tekin, Y. B.; Sentürk, S.; Sahin, F. K.; Avşar, A. F. Plasma Lipids and Lipoproteins during Pregnancy and Related Pregnancy Outcomes. *Archives of Gynecology and Obstetrics* **2013**, *288*, 49–55.

Engelke, U. F. H.; Tangerman, A.; Willemsen, M. A. A. P.; Moskau, D.; Loss, S.; Mudd, S. H.; Wevers, R. A. Dimethyl Sulfone in Human Cerebrospinal Fluid and Blood Plasma Confirmed by One-Dimensional ^{1}H and Two-Dimensional ^{1}H-^{13}C NMR. *NMR in Biomedicine* **2005**, *18*, 331–336.

EUROCAT Working Group. EUROCAT prevalence tables http://www.eurocat-network.eu/AccessPrevalenceData/PrevalenceTables.

Everard, A.; Cani, P. D. Diabetes, Obesity and Gut Microbiota. *Best Practice & Research. Clinical Gastroenterology* **2013**, *27*, 73–83.

Fabiano, A.; Gazzolo, D.; Zimmermann, L. J. I.; Gavilanes, A. W. D.; Paolillo, P.; Fanos, V.; Caboni, P.; Barberini, L.; Noto, A.; Atzori, L. Metabolomic Analysis of Bronchoalveolar Lavage Fluid in Preterm Infants Complicated by Respiratory Distress Syndrome: Preliminary Results. *The Journal of Maternal-Fetal & Neonatal Medicine* **2011**, *24 Suppl 2*, 55–58.

Fahy, E.; Sud, M.; Cotter, D.; Subramaniam, S. LIPID MAPS Online Tools for Lipid Research. *Nucleic Acids Research* **2007**, *35*, W606–W612.

Fahy, E.; Subramaniam, S.; Murphy, R. C.; Nishijima, M.; Raetz, C. R. H.; Shimizu, T.; Spener, F.; van Meer, G.; Wakelam, M. J. O.; Dennis, E. A. Update of the LIPID MAPS Comprehensive Classification System for Lipids. *Journal of Lipid Research* **2009**, *50 Suppl*, S9–S14.

Fan, T. W.-M. Metabolite Profiling by One- and Two-Dimensional NMR Analysis of Complex Mixtures. *Progress in Nuclear Magnetic Resonance Spectroscopy* **1996**, *28*, 161–219.

Fanos, V.; Atzori, L.; Makarenko, K.; Melis, G. B.; Ferrazzi, E. Metabolomics Application in Maternal-Fetal Medicine. *BioMed Research International* **2013a**, *2013*, 720514.

Fanos, V.; Van den Anker, J.; Noto, A.; Mussap, M.; Atzori, L. Metabolomics in Neonatology: Fact or Fiction? *Seminars in Fetal & Neonatal Medicine* **2013b**, *18*, 3–12.

Fanos, V.; Locci, E.; Noto, A.; Lazzarotto, T.; Manzoni, P.; Atzori, L.; Lanari, M. Urinary Metabolomics in Newborns Infected by Human Cytomegalovirus: A Preliminary Investigation. *Early Human Development* **2013c**, *89 Suppl 1*, S58–S61.

Fanos, V.; Pintus, M. C.; Lussu, M.; Atzori, L.; Noto, A.; Stronati, M.; Guimaraes, H.; Marcialis, M. A.; Rocha, G.; Moretti, C.; et al. Urinary Metabolomics of Bronchopulmonary Dysplasia (BPD): Preliminary Data at Birth Suggest It Is a Congenital Disease. *The Journal of Maternal-Fetal & Neonatal Medicine* **2014**, *27 Suppl 2*, 39–45.

Favretto, D.; Cosmi, E.; Ragazzi, E.; Visentin, S.; Tucci, M.; Fais, P.; Cecchetto, G.; Zanardo, V.; Viel, G.; Ferrara, S. D. Cord Blood Metabolomic Profiling in Intrauterine Growth Restriction. *Analytical and Bioanalytical Chemistry* **2012**, *402*, 1109–1121.

Fawcett, T. An Introduction to ROC Analysis. *Pattern Recogn. Lett.* **2006**, *27*, 861–874.

Fliniaux, O.; Gaillard, G.; Lion, A.; Cailleu, D.; Mesnard, F.; Betsou, F. Influence of Common Preanalytical Variations on the Metabolic Profile of Serum Samples in Biobanks. *Journal of Biomolecular NMR* **2011**, *51*, 457–465.

Gallos, I. D.; Sivakumar, K.; Kilby, M. D.; Coomarasamy, A.; Thangaratinam, S.; Vatish, M. Pre-Eclampsia Is Associated With, and Preceded By, Hypertriglyceridaemia: A Meta-Analysis. *BJOG: An International Journal of Obstetrics and Gynaecology* **2013**, *120*, 1321–1332.

Genc, H.; Uzun, H.; Benian, A.; Simsek, G.; Gelisgen, R.; Madazli, R.; Güralp, O. Evaluation of Oxidative Stress Markers in First Trimester for Assessment of Preeclampsia Risk. *Archives of Gynecology and Obstetrics* **2011**, *284*, 1367–1373.

George, F. H. M. *Diagnóstico E Conduta Na Diabetes Gestacional*; Norma da Direcção-Geral da Saúde N^{o} 007/2011: Lisboa, 2011.

Gil, A. M.; Duarte, I.; Cabrita, E.; Goodfellow, B. J.; Spraul, M.; Kerssebaum, R. Exploratory Applications of Diffusion Ordered Spectroscopy to Liquid Foods: An Aid towards Spectral Assignment. *Analytica Chimica Acta* **2004**, *506*, 215–223.

Di Giulio, A. M.; Carelli, S.; Castoldi, R. E.; Gorio, A.; Taricco, E.; Cetin, I. Plasma Amino Acid Concentrations throughout Normal Pregnancy and Early Stages of Intrauterine Growth Restricted Pregnancy. *Journal of Maternal-Fetal and Neonatal Medicine* **2004**, *15*, 356–362.

Glaser, C.; Demmelmair, H.; Koletzko, B. High-Throughput Analysis of Total Plasma Fatty Acid Composition with Direct in Situ Transesterification. *PLoS ONE* **2010**, *5*, e12045.

Golotvin, S.; Williams, A. Improved Baseline Recognition and Modeling of FT NMR Spectra. *Journal of Magnetic Resonance* **2000**, *146*, 122–125.

Gonzalez-Covarrubias, V.; Dane, A.; Hankemeier, T.; Vreeken, R. The Influence of Citrate, EDTA, and Heparin Anticoagulants to Human Plasma LC–MS Lipidomic Profiling. *Metabolomics* **2013**, *9*, 337–348.

Graça, G.; Duarte, I. F.; Barros, A. S.; Goodfellow, B. J.; Diaz, S.; Carreira, I. M.; Couceiro, A. B.; Galhano, E.; Gil, A. M. ^{1}H NMR Based Metabonomics of Human Amniotic Fluid for the Metabolic Characterization of Fetus Malformations. *Journal of Proteome Research* **2009**, *8*, 4144–4150.

Graça, G.; Duarte, I. F.; Barros, A. S.; Goodfellow, B. J.; Diaz, S. O.; Pinto, J.; Carreira, I. M.; Galhano, E.; Pita, C.; Gil, A. M.; et al. Impact of Prenatal Disorders on the Metabolic Profile of Second Trimester Amniotic Fluid: A Nuclear Magnetic Resonance Metabonomic Study. *Journal of Proteome Research* **2010**, *9*, 6016–6024.

Graça, G.; Diaz, S. O.; Pinto, J.; Barros, A. S.; Duarte, I. F.; Goodfellow, B. J.; Galhano, E.; Pita, C.; Almeida, M. do C.; Carreira, I. M.; et al. Can Biofluids Metabolic Profiling Help to Improve Healthcare during Pregnancy? *Spectroscopy: An International Journal* **2012a**, *27*, 515–523.

Graça, G.; Goodfellow, B. J.; Barros, A. S. A. S.; Diaz, S. S.; Duarte, I. F.; Spagou, K.; Veselkov, K.; Want, E. J.; Lindon, J. C.; Carreira, I. M.; et al. UPLC-MS Metabolic Profiling of Second Trimester Amniotic Fluid and Maternal Urine and Comparison with NMR Spectral Profiling for the Identification of Pregnancy Disorder Biomarkers. *Molecular BioSystems* **2012b**, *8*, 1243–1254.

Graça, G.; Moreira, A. S.; Correia, A. J. V; Goodfellow, B. J.; Barros, A. S.; Duarte, I. F.; Carreira, I. M.; Galhano, E.; Pita, C.; Almeida, M. do C.; et al. Mid-Infrared (MIR) Metabolic Fingerprinting of Amniotic Fluid: A Possible Avenue for Early Diagnosis of Prenatal Disorders? *Analytica Chimica Acta* **2013**, *764*, 24–31.

Gracie, S.; Pennell, C.; Ekman-Ordeberg, G.; Lye, S.; McManaman, J.; Williams, S.; Palmer, L.; Kelley, M.; Menon, R.; Gravett, M. An Integrated Systems Biology Approach to the Study of Preterm Birth Using “-Omic” Technology--a Guideline for Research. *BMC Pregnancy and Childbirth* **2011**, *11*, 71.

Griffiths, W. J.; Wang, Y. Mass Spectrometry: From Proteomics to Metabolomics and Lipidomics. *Chemical Society Reviews* **2009**, *38*, 1882–1896.

Griffiths, W. J.; Koal, T.; Wang, Y. Q.; Kohl, M.; Enot, D. P.; Deigner, H. P. Targeted Metabolomics for Biomarker Discovery. *Angewandte Chemie-International Edition* **2010**, *49*, 5426–5445.

Groenen, P. M. W.; Engelke, U. F.; Wevers, R. A.; Hendriks, J. C. M.; Eskes, T. K. A. B.; Merkus, H. M. W. M.; Steegers-Theunissen, R. P. M. High-Resolution ^{1}H NMR Spectroscopy of Amniotic Fluids from Spina Bifida Fetuses and Controls. *European Journal of Obstetrics & Gynecology and Reproductive Biology* **2004**, *112*, 16–23.

Gromski, P. S.; Xu, Y.; Correa, E.; Ellis, D. I.; Turner, M. L.; Goodacre, R. A Comparative Investigation of Modern Feature Selection and Classification Approaches for the Analysis of Mass Spectrometry Data. *Analytica Chimica Acta* **2014**, *829*, 1–8.

Gu, H.; Chen, H.; Pan, Z.; Jackson, A. U.; Talaty, N.; Xi, B.; Kissinger, C.; Duda, C.; Mann, D.; Raftery, D.; et al. Monitoring Diet Effects via Biofluids and Their Implications for Metabolomics Studies. *Analytical Chemistry* **2007**, *79*, 89–97.

Hacker, N. F.; Moore, J. G.; Gambone, J. C. *Essentials of Obstetrics and Gynecology.*, 4th ed.; Elsevier, 2004.

Hadden, D. R.; McLaughlin, C. Normal and Abnormal Maternal Metabolism during Pregnancy. *Seminars in Fetal and Neonatal Medicine* **2009**, *14*, 66–71.

Hahnemann, J. M.; Vejerslev, L. O. Accuracy of Cytogenetic Findings on Chorionic Villus Sampling (CVS)-Diagnostic Consequences of CVS Mosaicism and Non-Mosaic Discrepancy in Centres Contributing to EUCROMIC 1986-1992. *Prenatal Diagnosis* **1997**, *17*, 801–820.

Hammad, S. M.; Pierce, J. S.; Soodavar, F.; Smith, K. J.; Al Gadban, M. M.; Rembiesa, B.; Klein, R. L.; Hannun, Y. A.; Bielawski, J.; Bielawska, A. Blood Sphingolipidomics in Healthy Humans: Impact of Sample Collection Methodology. *Journal of Lipid Research* **2010**, *51*, 3074–3087.

Hannah, M. E.; Ohlsson, A.; Farine, D.; Hewson, S. A.; Hodnett, E. D.; Myhr, T. L.; Wang, E. E.; Weston, J. A.; Willan, A. R. Induction of Labor Compared with Expectant Management for Prelabor Rupture of the Membranes at Term. TERMPROM Study Group. *The New England Journal of Medicine* **1996**, *334*, 1005–1010.

Harkewicz, R.; Dennis, E. A. Applications of Mass Spectrometry to Lipids and Membranes. *Annual Review of Biochemistry* **2011**, *80*, 301–325.

Hashimoto, F.; Nishiumi, S.; Miyake, O.; Takeichi, H.; Chitose, M.; Ohtsubo, H.; Ishimori, S.; Ninchoji, T.; Hashimura, Y.; Kaito, H.; et al. Metabolomics Analysis of Umbilical Cord Blood Clarifies Changes in Saccharides Associated with Delivery Method. *Early Human Development* **2013**, *89*, 315–320.

Heazell, A. E. P.; Brown, M.; Dunn, W. B.; Worton, S. A.; Crocker, I. P.; Baker, P. N.; Kell, D. B. Analysis of the Metabolic Footprint and Tissue Metabolome of Placental Villous Explants Cultured at Different Oxygen Tensions Reveals Novel Redox Biomarkers. *Placenta* **2008**, *29*, 691–698.

Heazell, A. E. P.; Brown, M.; Worton, S. A.; Dunn, W. B. Review: The Effects of Oxygen on Normal and Pre-Eclamptic Placental Tissue--Insights from Metabolomics. *Placenta* **2011**, *32 Suppl 2*, S119–S124.

Heazell, A. E. P.; Bernatavicius, G.; Warrander, L.; Brown, M. C.; Dunn, W. B. A Metabolomic Approach Identifies Differences in Maternal Serum in Third Trimester Pregnancies That End in Poor Perinatal Outcome. *Reproductive Sciences* **2012**, *19*, 863–875.

Hebels, D. G. A. J.; Georgiadis, P.; Keun, H. C.; Athersuch, T. J.; Vineis, P.; Vermeulen, R.; Portengen, L.; Bergdahl, I. A.; Hallmans, G.; Palli, D.; et al. Performance in Omics Analyses of Blood Samples in Long-Term Storage: Opportunities for the Exploitation of Existing Biobanks in Environmental Health Research. *Environmental Health Perspectives* **2013**, *121*, 480–487.

Herrera, E.; Ortega, H. Metabolism in Normal Pregnancy. In *Textbook of Diabetes and Pregnancy*; Hod M., Jovanovic L., Di Renzo G. C., Leiva A., Langer O., Eds.; Informa Healthcare: London, 2008; pp 25–34.

Higby, K.; Suiter, C. R.; Phelps, J. Y.; Siler-Khodr, T.; Langer, O. Normal Values of Urinary Albumin and Total Protein Excretion during Pregnancy. *American Journal of Obstetrics and Gynecology* **1994**, *171*, 984–989.

Hoffmann, E. de; Stroobant, V. *Mass Spectrometry: Principles and Applications*, 3rd ed.; John Wiley & Sons Ltd: New Delhi, 2007.

Holmes, E.; Loo, R. L.; Cloarec, O.; Coen, M.; Tang, H.; Maibaum, E.; Bruce, S.; Chan, Q.; Elliott, P.; Stamler, J.; et al. Detection of Urinary Drug Metabolite (xenometabolome) Signatures in Molecular Epidemiology Studies via Statistical Total Correlation (NMR) Spectroscopy. *Analytical chemistry* **2007**, *79*, 2629–2640.

Honda, M.; Omori, Y.; Minei, S.; Oshiyama, T.; Shimizu, M.; Sanaka, M.; Kohama, T.; Nakabayashi, M.; Hirata, Y. Quantitative Analysis of Serum α1-Acid Glycoprotein Levels in Normal and Diabetic Pregnancy. *Diabetes Research and Clinical Practice* **1990**, *10*, 147–152.

Hore, P. J. NMR Principles. In *Encyclopedia of Spectroscopy and Spectrometry*; Lindon, J. C., Tranter, G. E., Holmes, J. L., Eds.; Elsevier, 1999; pp 1545–1553.

Horgan, R. P.; Clancy, O. H.; Myers, J. E.; Baker, P. N. An Overview of Proteomic and Metabolomic Technologies and Their Application to Pregnancy Research. *BJOG: An International Journal of Obstetrics and Gynaecology* **2009**, *116*, 173–181.

Horgan, R. P.; Broadhurst, D. I.; Dunn, W. B.; Brown, M.; Heazell, A. E. P.; Kell, D. B.; Baker, P. N.; Kenny, L. C. Changes in the Metabolic Footprint of Placental Explant-Conditioned Medium Cultured in Different Oxygen Tensions from Placentas of Small for Gestational Age and Normal Pregnancies. *Placenta* **2010**, *31*, 893–901.

Horgan, R. P.; Broadhurst, D. I.; Walsh, S. K.; Dunn, W. B.; Brown, M.; Roberts, C. T.; North, R. A.; McCowan, L. M.; Kell, D. B.; Baker, P. N.; et al. Metabolic Profiling Uncovers a Phenotypic Signature of Small for Gestational Age in Early Pregnancy. *Journal of Proteome Research* **2011**, *10*, 3660–3673.

Hübschmann, H.-J. *Handbook of GC/MS: Fundamentals and Applications*, 2nd ed.; WILEY-VCH Verlag GmbH & Co. KGaA: Weinheim, 2009.

Huynh, J.; Xiong, G.; Bentley-Lewis, R. A Systematic Review of Metabolite Profiling in Gestational Diabetes Mellitus. *Diabetologia* **2014**, *57*, 2453–2464.

Ilcol, Y. O.; Uncu, G.; Ulus, I. H.; Ozarda Ilcol, Y.; Uncu, G.; Ulus, I. H. Free and Phospholipid-Bound Choline Concentrations in Serum during Pregnancy, after Delivery and in Newborns. *Archives of Physiology and Biochemistry* **2002**, *110*, 393–399.

Issaq, H. J.; Van, Q. N.; Waybright, T. J.; Muschik, G. M.; Veenstra, T. D. Analytical and Statistical Approaches to Metabolomics Research. *Journal of Separation Science* **2009**, *32*, 2183–2199.

Ivorra, C.; Garcia-Vicent, C.; Chaves, F. J.; Monleon, D.; Manuel Morales, J.; Lurbe, E. Metabolomic Profiling in Blood from Umbilical Cords of Low Birth Weight Newborns. *Journal of Translational Medicine* **2012**, *10*, 1–10.

Jeyarajah, E. J.; Cromwell, W. C.; Otvos, J. D. Lipoprotein Particle Analysis by Nuclear Magnetic Resonance Spectroscopy. *Clinics in Laboratory Medicine* **2006**, *26*, 847–870.

Jiang, Z.; Liang, Q.; Wang, Y.; Zheng, X.; Pei, L.; Zhang, T.; Wang, Y.; Luo, G. Metabonomic Study on Women of Reproductive Age Treated with Nutritional Intervention: Screening Potential Biomarkers Related to Neural Tube Defects Occurrence. *Biomedical Chromatography* **2011**, *25*, 767–774.

Jolliffe, I. T. *Principal Component Analysis*, 2nd ed.; Springer Series in Statistics, 2002.

Kamath-Rayne, B. D.; Smith, H. C.; Muglia, L. J.; Morrow, A. L. Amniotic Fluid: The Use of High-Dimensional Biology to Understand Fetal Well-Being. *Reproductive Sciences* **2014**, *21*, 6–19.

Kane, S. C.; Costa, F. da S.; Brennecke, S. First Trimester Biomarkers in the Prediction of Later Pregnancy Complications. *BioMed Research International* **2014**, *2014*, 807196.

Katajamaa, M.; Orešič, M. Data Processing for Mass Spectrometry-Based Metabolomics. *Journal of Chromatography A* **2007**, *1158*, 318–328.

Keeler, J. Chapter 7. Two-Dimensional NMR. In *Understanding NMR spectroscopy*; Wiley, 2002a; pp 1–30.

Keeler, J. *Understanding NMR Spectroscopy*, 1st ed.; Wiley: Cambridge, 2002b.

Keeler, J.; Clowes, R. T.; Davis, A. L.; Laue, E. D. Pulsed-Field Gradients: Theory and Practice. *Methods in Enzymology* **1994**, *239*, 145–207.

Keeling, J. W.; Khong, T. Y. *Fetal and Neonatal Pathology*, 4th .; Spinger: London, 2007.

Kenny, L. C.; Broadhurst, D.; Brown, M.; Dunn, W. B.; Redman, C. W. G.; Kell, D. B.; Baker, P. N. Detection and Identification of Novel Metabolomic Biomarkers in Preeclampsia. *Reproductive Sciences* **2008**, *15*, 591–597.

Kenny, L. C.; Broadhurst, D. I.; Dunn, W.; Brown, M.; North, R. A.; McCowan, L.; Roberts, C.; Cooper, G. J. S.; Kell, D. B.; Baker, P. N. Robust Early Pregnancy Prediction of Later Preeclampsia Using Metabolomic Biomarkers. *Hypertension* **2010**, *56*, 741–749.

Kenny, L.; Dunn, W.; Ellis, D.; Myers, J.; Baker, P.; Consortium, G.; Kell, D. Novel Biomarkers for Pre-Eclampsia Detected Using Metabolomics and Machine Learning. *Metabolomics* **2005**, *1*, 227–234.

Kerssebaum, R. *DOSY and Diffusion by NMR*; Bruker BioSpin GmbH: Rheinstetten, 2006.

King, J. C. Physiology of Pregnancy and Nutrient Metabolism. *American Journal of Clinical Nutrition* **2000**, *71*, 1218S – 1225S.

Komarova, T. V; Petrunia, I. V; Shindyapina, A. V; Silachev, D. N.; Sheshukova, E. V; Kiryanov, G. I.; Dorokhov, Y. L. Endogenous Methanol Regulates Mammalian Gene Activity. *PLoS ONE* **2014**, *9*, e90239.

Korkes, H. A.; Sass, N.; Moron, A. F.; Câmara, N. O. S.; Bonetti, T.; Cerdeira, A. S.; Da Silva, I. D. C. G.; De Oliveira, L. Lipidomic Assessment of Plasma and Placenta of Women with Early-Onset Preeclampsia. *PLoS ONE* **2014**, *9*, e110747.

Koulman, A.; Prentice, P.; Wong, M. C. Y.; Matthews, L.; Bond, N. J.; Eiden, M.; Griffin, J. L.; Dunger, D. B. The Development and Validation of a Fast and Robust Dried Blood Spot Based Lipid Profiling Method to Study Infant Metabolism. *Metabolomics* **2014**, *10*, 1018–1025.

Kuc, S.; Koster, M. P. H.; Pennings, J. L. A.; Hankemeier, T.; Berger, R.; Harms, A. C.; Dane, A. D.; Schielen, P. C. J. I.; Visser, G. H. A.; Vreeken, R. J. Metabolomics Profiling for Identification of Novel Potential Markers in Early Prediction of Preeclampsia. *PLoS ONE* **2014**, *9*, e98540.

Kulig, C. C.; Beresford, T. P.; Everson, G. T. Rapid, Accurate, and Sensitive Fatty Acid Ethyl Ester Determination by Gas Chromatography-Mass Spectrometry. *Journal of Laboratory and Clinical Medicine* **2006**, *147*, 133–138.

Lämmerhofer, M.; Weckwerth, W. *Metabolomics in Practice: Successful Strategies to Generate and Analyze Metabolic Data*, 1st ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, 2013.

Lan, W.; Zhu, H.; Zhou, Z.; Ye, C.; Liu, M. ^{1}H NMR Investigation on Interaction between Ibuprofen and Lipoproteins. *Chemistry and Physics of Lipids* **2007**, *148*, 105–111.

Larsson, A.; Palm, M.; Hansson, L.-O.; Basu, S.; Axelsson, O. Reference Values for alpha1-Acid Glycoprotein, alpha1-Antitrypsin, Albumin, Haptoglobin, C-Reactive Protein, IgA, IgG and IgM during Pregnancy. *Acta Obstetricia et Gynecologica Scandinavica* **2008**, *87*, 1084–1088.

Lawton, K. A.; Berger, A.; Mitchell, M.; Milgram, K. E.; Evans, A. M.; Guo, L.; Hanson, R. W.; Kalhan, S. C.; Ryals, J. A.; Milburn, M. V. Analysis of the Adult Human Plasma Metabolome. *Pharmacogenomics* **2008**, *9*, 383–397.

Lee, D. Y.; Bowen, B. P.; Northen, T. R. Mass Spectrometry-Based Metabolomics, Analysis of Metabolite-Protein Interactions, and Imaging. *BioTechniques* **2010**, *49*, 557–565.

Levitan, I.; Volkov, S.; Subbaiah, P. V. Oxidized LDL: Diversity, Patterns of Recognition, and Pathophysiology. *Antioxidants & Redox Signaling* **2010**, *13*, 39–75.

Li, M.; Yang, L.; Bai, Y.; Liu, H. Analytical Methods in Lipidomics and Their Applications. *Analytical Chemistry* **2014**, *86*, 161–175.

Liang, X. P.; Wang, Y.; Liang, Q. L.; Wang, Y. M.; Huang, M.; Luo, G. A. Pathogenesis of Neural Tube Defects: The Story beyond Methylation or One-Carbon Unit Metabolism. *Metabolomics* **2012**, *8*, 919–929.

Liland, K. H. Multivariate Methods in Metabolomics – from Pre-Processing to Dimension Reduction and Statistical Analysis. *TrAC Trends in Analytical Chemistry* **2011**, *30*, 827–841.

Lin, X.; Yang, F.; Zhou, L.; Yin, P.; Kong, H.; Xing, W.; Lu, X.; Jia, L.; Wang, Q.; Xu, G. A Support Vector Machine-Recursive Feature Elimination Feature Selection Method Based on Artificial Contrast Variables and Mutual Information. *Journal of chromatography B* **2012**, *910*, 149–155.

Lindon, J. C.; Nicholson, J. K. Spectroscopic and Statistical Techniques for Information Recovery in Metabonomics and Metabolomics. *Annual Review of Analytical Chemistry* **2008**, *1*, 45–69.

Lindon, J. C.; Holmes, E.; Nicholson, J. K. Pattern Recognition Methods and Applications in Biomedical Magnetic Resonance. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2001**, *39*, 1–40.

Lindon, J. C.; Holmes, E.; Nicholson, J. K. Metabonomics and Its Role in Drug Development and Disease Diagnosis. *Expert Review of Molecular Diagnostics* **2004**, *4*, 189–199.

Lindon, J. C.; Nicholson, J. K.; Holmes, E. *Handbook of Metabonomics and Metabolomics*, 1st ed.; Elsevier: Amsterdam ; Boston, 2007.

Lippi, G.; Albiero, A.; Montagnana, M.; Salvagno, G. L.; Scevarolli, S.; Franchi, M.; Guidi, G. C. Lipid and Lipoprotein Profile in Physiological Pregnancy. *Clinical Laboratory* **2007**, *53*, 173–177.

Liu, H.; Zhang, Z.; Linhardt, R. J. Lessons Learned from the Contamination of Heparin. *Natural Product Reports* **2009**, *26*, 313–321.

Liu, L.; Li, Y.; Guan, C.; Li, K.; Wang, C.; Feng, R.; Sun, C. Free Fatty Acid Metabolic Profile and Biomarkers of Isolated Post-Challenge Diabetes and Type 2 Diabetes Mellitus Based on GC–MS and Multivariate Statistical Analysis. *Journal of Chromatography B* **2010**, *878*, 2817–2825.

Liu, M. L.; Tang, H. R.; Nicholson, J. K.; Lindon, J. C. Use of ^{1}H NMR-Determined Diffusion Coefficients to Characterize Lipoprotein Fractions in Human Blood Plasma. *Magnetic Resonance in Chemistry* **2002**, *40*, S83–S88.

Longini, M.; Tataranno, M. L.; Proietti, F.; Tortoriello, M.; Belvisi, E.; Vivi, A.; Tassini, M.; Perrone, S.; Buonocore, G. A Metabolomic Study of Preterm and Term Human and Formula Milk by Proton MRS Analysis: Preliminary Results. *The Journal of Maternal-Fetal & Neonatal Medicine* **2014**, *27 Suppl 2*, 27–33.

Lowe, W. L.; Karban, J. Genetics, Genomics and Metabolomics: New Insights into Maternal Metabolism during Pregnancy. *Diabetic Medicine* **2014**, *31*, 254–262.

Lu, X.; Ye, M. *Enrichment of Phospholipids in Biological Samples Using HybridSPE-PL*; 2010.

Luan, H.; Meng, N.; Liu, P.; Feng, Q.; Lin, S.; Fu, J.; Davidson, R.; Chen, X.; Rao, W.; Chen, F.; et al. Pregnancy-Induced Metabolic Phenotype Variations in Maternal Plasma. *Journal of Proteome Research* **2014**, *13*, 1527–1536.

Mahadevan, S.; Shah, S. L.; Marrie, T. J.; Slupsky, C. M. Analysis of Metabolomic Data Using Support Vector Machines. *Analytical Chemistry* **2008**, *80*, 7562–7570.

Maher, A. D.; Cysique, L. A.; Brew, B. J.; Rae, C. D. Statistical Integration of ^{1}H NMR and MRS Data from Different Biofluids and Tissues Enhances Recovery of Biological Information from Individuals with HIV-1 Infection. *Journal of Proteome Research* **2011**, *10*, 1737–1745.

Mallol, R.; Rodriguez, M. A.; Brezmes, J.; Masana, L.; Correig, X. Human Serum/plasma Lipoprotein Analysis by NMR: Application to the Study of Diabetic Dyslipidemia. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2013**, *70*, 1–24.

Mankuta, D.; Elami-Suzin, M.; Elhayani, A.; Vinker, S. Lipid Profile in Consecutive Pregnancies. *Lipids in Health and Disease* **2010**, *9*, 58.

Marincola, F. C.; Noto, A.; Caboni, P.; Reali, A.; Barberini, L.; Lussu, M.; Murgia, F.; Santoru, M. L.; Atzori, L.; Fanos, V. A Metabolomic Study of Preterm Human and Formula Milk by High Resolution NMR and GC/MS Analysis: Preliminary Results. *The Journal of Maternal-Fetal & Neonatal Medicine* **2012**, *25*, 62–67.

Matyash, V.; Liebisch, G.; Kurzchalia, T. V; Shevchenko, A.; Schwudke, D. Lipid Extraction by Methyl-Tert-Butyl Ether for High-Throughput Lipidomics. *Journal of Lipid Research* **2008**, *49*, 1137–1146.

Meiboom, S.; Gill, D. Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. *Review of Scientific Instruments* **1958**, *29*, 688.

Menon, R.; Jones, J.; Gunst, P. R.; Kacerovsky, M.; Fortunato, S. J.; Saade, G. R.; Basraon, S. Amniotic Fluid Metabolomic Analysis in Spontaneous Preterm Birth. *Reproductive Sciences* **2014**, *21*, 791–803.

Metzger, B. E.; Gabbe, S. G.; Persson, B.; Buchanan, T. A.; Catalano, P. A.; Damm, P.; Dyer, A. R.; Leiva, A. de; Hod, M.; Kitzmiler, J. L.; et al. International Association of Diabetes and Pregnancy Study Groups Recommendations on the Diagnosis and Classification of Hyperglycemia in Pregnancy. *Diabetes Care* **2010**, *33*, 676–682.

Mishur, R. J.; Rea, S. L. Applications of Mass Spectrometry to Metabolomics and Metabonomics: Detection of Biomarkers of Aging and of Age-Related Diseases. *Mass Spectrometry Reviews* **2011**, *31*, 70–95.

Moco, S.; Collino, S.; Rezzi, S.; Martin, F.-P. J. Metabolomics Perspectives in Pediatric Research. *Pediatric Research* **2013**, *73*, 570–576.

Moltu, S. J.; Sachse, D.; Blakstad, E. W.; Strømmen, K.; Nakstad, B.; Almaas, A. N.; Westerberg, A. C.; Rønnestad, A.; Brække, K.; Veierød, M. B.; et al. Urinary Metabolite Profiles in Premature Infants Show Early Postnatal Metabolic Adaptation and Maturation. *Nutrients* **2014**, *6*, 1913–1930.

Moore, K. L.; Persaud, T. V. N. *The Developing Human: Clinically Oriented Embryology*; Elsevier Saunders: Philadelphia, PA, 2008.

Moore, R. M.; Mansour, J. M.; Redline, R. W.; Mercer, B. M.; Moore, J. J. The Physiology of Fetal Membrane Rupture: Insight Gained from the Determination of Physical Properties. *Placenta* **2006**, *27*, 1037–1051.

Mudd, L. M.; Holzman, C. B.; Catov, J. M.; Senagore, P. K.; Evans, R. W. Maternal Lipids at Mid-Pregnancy and the Risk of Preterm Delivery. *Acta Obstetricia et Gynecologica Scandinavica* **2012**, *91*, 726–735.

Mussap, M.; Antonucci, R.; Noto, A.; Fanos, V. The Role of Metabolomics in Neonatal and Pediatric Laboratory Medicine. *Clinica Chimica Acta* **2013**, *426*, 127–138.

Nelson, D. L.; Cox, M. M. *Lehninger Principles of Biochemistry*, 4th ed.; W H Freeman & Co Publication , 2005.

Nicholson, J. K.; Foxall, P. J. D.; Spraul, M.; Farrant, R. D.; Lindon, J. C. 750-MHz ^{1}H and ^{1}H-^{13}C NMR Spectroscopy of Human Blood Plasma. *Analytical Chemistry* **1995**, *67*, 793–811.

Nicholson, J. K.; Lindon, J. C.; Holmes, E. "Metabonomics": Understanding the Metabolic Responses of Living Systems to Pathophysiological Stimuli via Multivariate Statistical Analysis of Biological NMR Spectroscopic Data. *Xenobiotica* **1999**, *29*, 1181–1189.

Nicholson, J. K.; Connelly, J.; Lindon, J. C.; Holmes, E. Metabonomics: A Platform for Studying Drug Toxicity and Gene Function. *Nature Reviews Drug Discovery* **2002**, *1*, 153–161.

Nicolaides, K. H. Nuchal Translucency and Other First-Trimester Sonographic Markers of Chromosomal Abnormalities. *American Journal of Obstetrics and Gynecology* **2004**, *191*, 45–67.

Nicolaides, K. H. Screening for Fetal Aneuploidies at 11 to 13 Weeks. *Prenatal Diagnosis* **2011**, *31*, 7–15.

Nishiumi, S.; Kobayashi, T.; Ikeda, A.; Yoshie, T.; Kibi, M.; Izumi, Y.; Okuno, T.; Hayashi, N.; Kawano, S.; Takenawa, T.; et al. A Novel Serum Metabolomics-Based Diagnostic Approach for Colorectal Cancer. *PLoS ONE* **2012**, *7*, e40459.

Odibo, A. O.; Goetzinger, K. R.; Odibo, L.; Cahill, A. G.; Macones, G. A.; Nelson, D. M.; Dietzen, D. J. First-Trimester Prediction of Preeclampsia Using Metabolomic Biomarkers: A Discovery Phase Study. *Prenatal Diagnosis* **2011**, *31*, 990–994.

De Oliveira, L.; Câmara, N. O. S.; Bonetti, T.; Turco, E. G. Lo; Bertolla, R. P.; Moron, A. F.; Sass, N.; Da Silva, I. D. C. G. *Lipid Fingerprinting in Women with Early-Onset Preeclampsia: A First Look*; 2012; Vol. 45.

Oostendorp, M.; Engelke, U. F. H.; Willemsen, M. A. A. P.; Wevers, R. A. Diagnosing Inborn Errors of Lipid Metabolism with Proton Nuclear Magnetic Resonance Spectroscopy. *Clinical Chemistry* **2006**, *52*, 1395–1405.

Ottolenghi, C.; Abermil, N.; Lescoat, A.; Aupetit, J.; Beaugendre, O.; Morichon-Delvallez, N.; Ricquier, D.; Chadefaux-Vekemans, B.; Rabier, D. Gestational Age-Related Reference Values for Amniotic Fluid Organic Acids. *Prenatal Diagnosis* **2010**, *30*, 43–48.

Otvos, J. D.; Jeyarajah, E. J.; Bennett, D. W. Quantification of Plasma Lipoproteins by Proton Nuclear Magnetic Resonance Spectroscopy. *Clinical Chemistry* **1991**, *37*, 377–386.

Otvos, J. D.; Jeyarajah, E. J.; Bennett, D. W.; Krauss, R. M. Development of a Proton Nuclear Magnetic Resonance Spectroscopic Method for Determining Plasma Lipoprotein Concentrations and Subspecies Distributions from a Single, Rapid Measurement. *Clinical Chemistry* **1992**, *38*, 1632–1638.

Pallotto, E. K.; Kilbride, H. W. Perinatal Outcome and Later Implications of Intrauterine Growth Restriction. *Clinical Obstetrics and Gynecology* **2006**, *49*, 257–269.

Park, Y.; Kim, S. B.; Wang, B.; Blanco, R. A.; Le, N.-A.; Wu, S.; Accardi, C. J.; Alexander, R. W.; Ziegler, T. R.; Jones, D. P. Individual Variation in Macronutrient Regulation Measured by Proton Magnetic Resonance Spectroscopy of Human Plasma. *American Journal of Physiology-Regulatory Integrative and Comparative Physiology* **2009**, *297*, R202–R209.

Parsons, H. M.; Ekman, D. R.; Collette, T. W.; Viant, M. R. Spectral Relative Standard Deviation: A Practical Benchmark in Metabolomics. *The Analyst* **2009**, *134*, 478–485.

Pasikanti, K. K.; Ho, P. C.; Chan, E. C. Y. Gas Chromatography/mass Spectrometry in Metabolic Profiling of Biological Fluids. *Journal of Chromatography B* **2008**, *871*, 202–211.

Peré-Trepat, E.; Ross, A. B.; Martin, F.-P.; Rezzi, S.; Kochhar, S.; Hasselbalch, A. L.; Kyvik, K. O.; Sørensen, T. I. A. Chemometric Strategies to Assess Metabonomic Imprinting of Food Habits in Epidemiological Studies. *Chemometrics and Intelligent Laboratory Systems* **2010**, *104*, 95–100.

Perluigi, M.; di Domenico, F.; Fiorini, A.; Cocciolo, A.; Giorgi, A.; Foppoli, C.; Butterfield, D. A.; Giorlandino, M.; Giorlandino, C.; Schininà, M. E.; et al. Oxidative Stress Occurs Early in Down Syndrome Pregnancy: A Redox Proteomics Analysis of Amniotic Fluid. *Proteomics. Clinical Applications* **2011**, *5*, 167–178.

Pinto, J.; Domingues, M. R.; Gil, A. M. Blood Metabolomics in Human Prenatal and Newborn Health Studies. In *Global Metabolic Profiling: Clinical Applications*; Nichols, A., Theodoridis, G., Wilson, I. D., Eds.; Future Science Ltd: London, 2014a; pp 50–68; eISBN: 978-1-909453-61-6.

Pinto, J.; Barros, A. S.; Domingues, M. R. M.; Goodfellow, B. J.; Galhano, E.; Pita, C.; Almeida, M. C.; Carreira, I. M.; Gil, A. M. Following Healthy Pregnancy by NMR Metabolomics of Plasma and Correlation to Urine. *Journal of Proteome Research* **2014b**, *14*, 1263–1274.

Pinto, J.; Domingues, M. R. M.; Galhano, E.; Pita, C.; Almeida, M. do C.; Carreira, I. M.; Gil, A. M. Human Plasma Stability during Handling and Storage: Impact on NMR Metabolomics. *The Analyst* **2014c**, *139*, 1168–1177.

Pinto, J.; Maciel, E.; Melo, T. S.; Domingues, M. R. M.; Galhano, E.; Pita, C.; Almeida, M. do C.; Carreira, I. M.; Gil, A. M. Maternal Plasma Phospholipids Are Altered in Trisomy 21 Cases and prior to Preeclampsia and Preterm Outcomes. *Rapid communications in Mass Spectrometry* **2014d**, *28*, 1635–1638.

Pluskal, T.; Castillo, S.; Villar-Briones, A.; Oresic, M. MZmine 2: Modular Framework for Processing, Visualizing, and Analyzing Mass Spectrometry-Based Molecular Profile Data. *BMC Bioinformatics* **2010**, *11*, 395.

Power, K. M.; Sanchez-Galan, J. E.; Luskey, G. W.; Koski, K. G.; Burns, D. H. Use of near-Infrared Spectroscopic Analysis of Second Trimester Amniotic Fluid to Assess Preterm Births. *Journal of Pregnancy* **2011**, *2011*, 980985.

Pridjian, G.; Benjamin, T. D. Update on Gestational Diabetes. *Obstetrics and Gynecology Clinics of North America* **2010**, *37*, 255–267.

Quehenberger, O.; Armando, A. M.; Brown, A. H.; Milne, S. B.; Myers, D. S.; Merrill, A. H.; Bandyopadhyay, S.; Jones, K. N.; Kelly, S.; Shaner, R. L.; et al. Lipidomics Reveals a Remarkable Diversity of Lipids in Human Plasma. *Journal of Lipid Research* **2010**, *51*, 3299–3305.

Quintás, G.; Portillo, N.; García-Cañaveras, J.; Castell, J.; Ferrer, A.; Lahoz, A. Chemometric Approaches to Improve PLSDA Model Outcome for Predicting Human Non-Alcoholic Fatty Liver Disease Using UPLC-MS as a Metabolic Profiling Tool. *Metabolomics* **2009**, *8*, 86–98.

R. D. C. Team. *R: A Language and Environment for Statistical Computing*, 2.14.1 ed.; R Foundation for Statistical Computing: Vienna, Austria, 2010.

Rajalahti, T.; Arneberg, R.; Berven, F. S.; Myhr, K.-M.; Ulvik, R. J.; Kvalheim, O. M. Biomarker Discovery in Mass Spectral Profiles by Means of Selectivity Ratio Plot. *Chemometrics and Intelligent Laboratory Systems* **2009**, *95*, 35–48.

Rankin, N. J.; Preiss, D.; Welsh, P.; Burgess, K. E. V; Nelson, S. M.; Lawlor, D. A.; Sattar, N. The Emergence of Proton Nuclear Magnetic Resonance Metabolomics in the Cardiovascular Arena as Viewed from a Clinical Perspective. *Atherosclerosis* **2014**, *237*, 287–300.

Raymond, D.; Peterson, E. A Critical Review of Early-Onset and Late-Onset Preeclampsia. *Obstetrical & Gynecological Survey* **2011**, *66*, 497–506.

Reddy, G. N. M.; Caldarelli, S. Demixing of Severely Overlapping NMR Spectra through Multiple-Quantum NMR. *Analytical Chemistry* **2010**, *82*, 3266–3269.

Reddy, G. N. M.; Caldarelli, S. Maximum-Quantum (MaxQ) NMR for the Speciation of Mixtures of Phenolic Molecules. *Chemical Communications (Cambridge, England)* **2011**, *47*, 4297–4299.

Reddy, G. N. M.; Caldarelli, S. Identification and Quantification of EPA 16 Priority Polycyclic Aromatic Hydrocarbon Pollutants by Maximum-Quantum NMR. *The Analyst* **2012**, *137*, 741–746.

Reinke, S. N.; Walsh, B. H.; Boylan, G. B.; Sykes, B. D.; Kenny, L. C.; Murray, D. M.; Broadhurst, D. I. ^{1}H NMR Derived Metabolomic Profile of Neonatal Asphyxia in Umbilical Cord Serum: Implications for Hypoxic Ischemic Encephalopathy. *Journal of Proteome Research* **2013**, *12*, 4230–4239.

Rezzi, S.; Collino, S.; Goulet, L.; Martin, F.-P. Metabonomic Approaches to Nutrient Metabolism and Future Molecular Nutrition. *TrAC Trends in Analytical Chemistry* **2013**, *52*, 112–119.

Rinnan, Å.; Andersson, M.; Ridder, C.; Engelsen, S. B. Recursive Weighted Partial Least Squares (rPLS): An Efficient Variable Selection Method Using PLS. *Journal of Chemometrics* **2014**, *28*, 439–447.

Rolim, A. E. H.; Henrique-Araújo, R.; Ferraz, E. G.; de Araújo Alves Dultra, F. K.; Fernandez, L. G. Lipidomics in the Study of Lipid Metabolism: Current Perspectives in the Omic Sciences. *Gene* **2014**, *554*, 131–139.

Romero, R.; Espinoza, J.; Gotsch, F.; Kusanovic, J. P.; Friel, L. A.; Erez, O.; Mazaki-Tovi, S.; Than, N. G.; Hassan, S.; Tromp, G. The Use of High-Dimensional Biology (genomics, Transcriptomics, Proteomics, and Metabolomics) to Understand the Preterm Parturition Syndrome. *BJOG: An International Journal of Obstetrics and Gynaecology* **2006**, *113 Suppl* , 118–135.

Romero, R.; Mazaki-Tovi, S.; Vaisbuch, E.; Kusanovic, J. P.; Chaiworapongsa, T.; Gomez, R.; Nien, J. K.; Yoon, B. H.; Mazor, M.; Luo, J. Q.; et al. Metabolomics in Premature Labor: A Novel Approach to Identify Patients at Risk for Preterm Delivery. *Journal of Maternal-Fetal & Neonatal Medicine* **2010**, *23*, 1344–1359.

Rousseau, R.; Govaerts, B.; Verleysen, M.; Boulanger, B. Comparison of Some Chemometric Tools for Metabonomics Biomarker Identification. *Chemometrics and Intelligent Laboratory Systems* **2008**, *91*, 54–66.

Sachse, D.; Sletner, L.; Mørkrid, K.; Jenum, A. K.; Birkeland, K. I.; Rise, F.; Piehler, A. P.; Berg, J. P. Metabolic Changes in Urine during and after Pregnancy in a Large, Multiethnic Population-Based Cohort Study of Gestational Diabetes. *PLoS ONE* **2012**, *7*, e52399.

Sachse, D.; Bærug, A.; Sletner, L.; Birkeland, K. I.; Nakstad, B.; Jenum, A. K.; Berg, J. P. Urine NMR Metabolomics Analysis of Breastfeeding Biomarkers during and after Pregnancy in a Large Prospective Cohort Study. *Scandinavian Journal of Clinical and Laboratory Investigation* **2014**, *74*, 264–272.

Sahoo, S.; Franzson, L.; Jonsson, J. J.; Thiele, I. A Compendium of Inborn Errors of Metabolism Mapped onto the Human Metabolic Network. *Molecular Biosystems* **2012**, *8*, 2545–2558.

Dos Santos-Weiss, I. C. R.; Réa, R. R.; Fadel-Picheth, C. M. T.; Rego, F. G. M.; Pedrosa, F. de O.; Gillery, P.; Souza, E. M.; Picheth, G. The Plasma Logarithm of the triglyceride/HDL-Cholesterol Ratio Is a Predictor of Low Risk Gestational Diabetes in Early Pregnancy. *Clinica Chimica Acta* **2013**, *418*, 1–4.

Sanz-Cortés, M.; Carbajo, R. J.; Crispi, F.; Figueras, F.; Pineda-Lucena, A.; Gratacós, E. Metabolomic Profile of Umbilical Cord Blood Plasma from Early and Late Intrauterine Growth Restricted (IUGR) Neonates with and without Signs of Brain Vasodilation. *PLoS ONE* **2013**, *8*, e80121.

Sarafian, M. H.; Gaudin, M.; Lewis, M. R.; Martin, F.-P.; Holmes, E.; Nicholson, J. K.; Dumas, M.-E. Objective Set of Criteria for Optimization of Sample Preparation Procedures for Ultra-High Throughput Untargeted Blood Plasma Lipid Profiling by Ultra Performance Liquid Chromatography-Mass Spectrometry. *Analytical Chemistry* **2014**, *86*, 5766–5774.

Scherer, M.; Leuthäuser-Jaschinski, K.; Ecker, J.; Schmitz, G.; Liebisch, G. A Rapid and Quantitative LC-MS/MS Method to Profile Sphingolipids. *Journal of Lipid Research* **2010**, *51*, 2001–2011.

Scherer, M.; Böttcher, A.; Schmitz, G.; Liebisch, G. Sphingolipid Profiling of Human Plasma and FPLC-Separated Lipoprotein Fractions by Hydrophilic Interaction Chromatography Tandem Mass Spectrometry. *Biochimica et Biophysica Acta* **2011**, *1811*, 68–75.

Scholtens, D. M.; Muehlbauer, M. J.; Daya, N. R.; Stevens, R. D.; Dyer, A. R.; Lowe, L. P.; Metzger, B. E.; Newgard, C. B.; Bain, J. R.; Lowe, W. L. Metabolomics Reveals Broad-Scale Metabolic Perturbations in Hyperglycemic Mothers during Pregnancy. *Diabetes Care* **2014**, *37*, 158–166.

Seavey, B. R.; Farr, E. A.; Westler, W. M.; Markley, J. L. A Relational Database for Sequence-Specific Protein NMR Data. *Journal of Biomolecular NMR* **1991**, *1*, 217–236.

De Seymour, J. V; Conlon, C. A.; Sulek, K.; Villas Bôas, S. G.; McCowan, L. M. E.; Kenny, L. C.; Baker, P. N. Early Pregnancy Metabolite Profiling Discovers a Potential Biomarker for the Subsequent Development of Gestational Diabetes Mellitus. *Acta Diabetologica* **2014**, *51*, 887–890.

Shah, S. H.; Kraus, W. E.; Newgard, C. B. Metabolomic Profiling for the Identification of Novel Biomarkers and Mechanisms Related to Common Cardiovascular Diseases: Form and Function. *Circulation* **2012**, *126*, 1110–1120.

Sheskin, D. J. Parametric Versus Nonparametric Tests. In *International Encyclopedia of Statistical Science*; Lovric, M., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2014; pp 1051–1052.

Sitter, B.; Sonnewald, U.; Spraul, M.; Fjösne, H. E.; Gribbestad, I. S. High-Resolution Magic Angle Spinning MRS of Breast Cancer Tissue. *NMR in Biomedicine* **2002**, *15*, 327–337.

Smith, L. M.; Maher, A. D.; Cloarec, O.; Rantalainen, M.; Tang, H.; Elliott, P.; Stamler, J.; Lindon, J. C.; Holmes, E.; Nicholson, J. K. Statistical Correlation and Projection Methods for Improved Information Recovery from Diffusion-Edited NMR Spectra of Biological Samples. *Analytical Chemistry* **2007**, *79*, 5682–5689.

Smolinska, A.; Blanchet, L.; Buydens, L. M. C.; Wijmenga, S. S. NMR and Pattern Recognition Methods in Metabolomics: From Data Acquisition to Biomarker Discovery: A Review. *Analytica Chimica Acta* **2012a**, *750*, 82–97.

Smolinska, A.; Posma, J. M.; Blanchet, L.; Ampt, K. A. M.; Attali, A.; Tuinstra, T.; Luider, T.; Doskocz, M.; Michiels, P. J.; Girard, F. C.; et al. Simultaneous Analysis of Plasma and CSF by NMR and Hierarchical Models Fusion. *Analytical and Bioanalytical Chemistry* **2012b**, *403*, 947–959.

Soininen, P.; Öörni, K.; Maaheimo, H.; Laatikainen, R.; Kovanen, P. T.; Kaski, K.; Ala-Korpela, M. ^{1}H NMR at 800 MHz Facilitates Detailed Phospholipid Follow-up during Atherogenic Modifications in Low Density Lipoproteins. *Biochemical and Biophysical Research Communications* **2007**, *360*, 290–294.

Spagou, K.; Tsoukali, H.; Raikos, N.; Gika, H.; Wilson, I. D.; Theodoridis, G. Hydrophilic Interaction Chromatography Coupled to MS for Metabonomic/metabolomic Studies. *Journal of Separation Science* **2010**, *33*, 716–727.

Srivastava, N. K.; Pradhan, S.; Mittal, B.; Gowda, G. A. N. High Resolution NMR Based Analysis of Serum Lipids in Duchenne Muscular Dystrophy Patients and Its Possible Diagnostic Significance. *NMR in Biomedicine* **2010**, *23*, 13–22.

Sulek, K.; Han, T.-L.; Villas-Boas, S. G.; Wishart, D. S.; Soh, S.-E.; Kwek, K.; Gluckman, P. D.; Chong, Y.-S.; Kenny, L. C.; Baker, P. N. Hair Metabolomics: Identification of Fetal Compromise Provides Proof of Concept for Biomarker Discovery. *Theranostics* **2014**, *4*, 953–959.

Sutcliffe, R. G. The Nature and Origin of the Soluble Protein in Human Amniotic Fluid. *Biological Reviews* **1975**, *50*, 1–33.

Syggelou, A.; Iacovidou, N.; Atzori, L.; Xanthos, T.; Fanos, V. Metabolomics in the Developing Human Being. *Pediatric Clinics of North America* **2012**, *59*, 1039–1058.

Tabor, A.; Alfirevic, Z. Update on Procedure-Related Risks for Prenatal Diagnosis Techniques. *Fetal Diagnosis and Therapy* **2010**, *27*, 1–7.

Tarim, E.; Bagis, T.; Kilicdag, E.; Erkanli, S.; Aslan, E.; Sezgin, N.; Kuscu, E. Elevated Plasma Homocysteine Levels in Gestational Diabetes Mellitus. *Acta Obstetricia et Gynecologica Scandinavica* **2004**, *83*, 543–547.

Tea, I.; Le Gall, G.; Küster, A.; Guignard, N.; Alexandre–Gouabau, M.-C.; Darmaun, D.; Robins, R. J. ^{1}H-NMR-Based Metabolic Profiling of Maternal and Umbilical Cord Blood Indicates Altered Materno-Foetal Nutrient Exchange in Preterm Infants. *PLoS ONE* **2012**, *7*, e29947.

Teahan, O.; Gamble, S.; Holmes, E.; Waxman, J.; Nicholson, J. K.; Bevan, C.; Keun, H. C. Impact of Analytical Bias in Metabonomic Studies of Human Blood Serum and Plasma. *Analytical Chemistry* **2006**, *78*, 4307–4318.

Than, N. G.; Vaisbuch, E.; Kim, C. J.; Mazaki-Tovi, S.; Erez, O.; Yeo, L.; Mittal, P.; Hupuczi, Petronella, Varkonyi, T.; Hassan, S. S.; Papp, Z.; et al. Early-Onset Preeclampsia and HELLP Syndrome: An Overview. In *Handbook of Growth and Growth Monitoring in Health and Disease*; Preedy, V. R., Ed.; Springer New York, 2012; pp 1867–1891.

Thorp, J. M.; Rice, M. M.; Harper, M.; Klebanoff, M.; Sorokin, Y.; Varner, M. W.; Wapner, R. J.; Caritis, S. N.; Iams, J. D.; Peaceman, A. M.; et al. Advanced Lipoprotein Measures and Recurrent Preterm Birth. *American Journal of Obstetrics and Gynecology* **2013**, *209*, 342.e1–e7.

Tongsong, T.; Wanapirak, C.; Kunavikatikul, C.; Sirirchotiyakul, S.; Piyamongkol, W.; Chanprapaph, P. Cordocentesis at 16-24 Weeks of Gestation: Experience of 1,320 Cases. *Prenatal Diagnosis* **2000**, *20*, 224–228.

Torri, G. M.; Torri, J.; Gulian, J.-M. M.; Vion-Dury, J.; Viout, P.; Cozzone, P. J.; J. Cozzone, P. Magnetic Resonance Spectroscopy of Serum and Acute-Phase Proteins Revisited: A Multiparametric Statistical Analysis of Metabolite Variations in Inflammatory, Infectious and Miscellaneous Diseases. *Clinica Chimica Acta* **1999**, *279*, 77–96.

Tovar, A. M. F.; Capillé, N. V. M.; Santos, G. R. C.; Vairo, B. C.; Oliveira, S.-N. M. C. G.; Fonseca, R. J. C.; Mourão, P. A. S. Heparin from Bovine Intestinal Mucosa: Glycans with Multiple Sulfation Patterns and Anticoagulant Effects. *Thrombosis and Haemostasis* **2012**, *107*, 903–915.

Trabi, M.; Keller, M.; Jonsson, N. NMR-Based Metabonomics of Bovine Blood: An Investigation into the Effects of Long Term Storage on Plasma Samples. *Metabolomics* **2013**, 1–7.

Triba, M. N.; Le Moyec, L.; Amathieu, R.; Goossens, C.; Bouchemal, N.; Nahon, P.; Rutledge, D. N.; Savarin, P. PLS/OPLS Models in Metabolomics: The Impact of Permutation of Dataset Rows on the K-Fold Cross-Validation Quality Parameters. *Molecular BioSystems* **2015**, *11*, 13–19.

Trivedi, D. K.; Iles, R. K. HILIC-MS-Based Shotgun Metabolomic Profiling of Maternal Urine at 9-23 Weeks of Gestation - Establishing the Baseline Changes in the Maternal Metabolome. *Biomedical Chromatography: BMC* **2014a**, *29*, 240–245.

Trivedi, D. K.; Iles, R. K. Shotgun Metabolomic Profiles in Maternal Urine Identify Potential Mass Spectral Markers of Abnormal Fetal Biochemistry - Dihydrouracil and Progesterone in the Metabolism of Down Syndrome. *Biomedical Chromatography* **2014b**, DOI: 10.1002/bmc.3404.

Trygg, J.; Holmes, E.; Lundstedt, T. Chemometrics in Metabonomics. *Journal of Proteome Research* **2007**, *6*, 469–479.

Tukiainen, T.; Tynkkynen, T.; Mäkinen, V.-P.; Jylänki, P.; Kangas, A.; Hokkanen, J.; Vehtari, A.; Gröhn, O.; Hallikainen, M.; Soininen, H.; et al. A Multi-Metabolite Analysis of Serum by ^{1}H NMR Spectroscopy: Early Systemic Signs of Alzheimer's Disease. *Biochemical and Biophysical Research Communications* **2008**, *375*, 356–361.

Turner, E.; Brewster, J. A.; Simpson, N. A. B.; Walker, J. J.; Fisher, J. Plasma from Women with Preeclampsia Has a Low Lipid and Ketone Body Content - A Nuclear Magnetic Resonance Study. *Hypertension in Pregnancy* **2007**, *26*, 329–342.

Turner, E.; Brewster, J. A.; Simpson, N. A. B.; Walker, J. J.; Fisher, J. Aromatic Amino Acid Biomarkers of Preeclampsia - A Nuclear Magnetic Resonance Investigation. *Hypertension in Pregnancy* **2008**, *27*, 225–235.

Uzan J; Carbonnel M; Piconne O; Asmar R; Ayoubi J-M. Pre-Eclampsia: Pathophysiology, Diagnosis, and Management. *Vascular Health and Risk Management* **2011**, *7*, 467–474.

Vaz, F. M.; Pras-Raves, M.; Bootsma, A. H.; van Kampen, A. H. C. Principles and Practice of Lipidomics. *Journal of Inherited Metabolic Disease* **2015**, *38*, 41–52.

Veselkov, K. A.; Lindon, J. C.; Ebbels, T. M. D.; Crockford, D.; Volynkin, V. V; Holmes, E.; Davies, D. B.; Nicholson, J. K. Recursive Segment-Wise Peak Alignment of Biological ^{1}H NMR Spectra for Improved Metabolic Biomarker Recovery. *Analytical Chemistry* **2009**, *81*, 56–66.

Veselkov, K. A.; Vingara, L. K.; Masson, P.; Robinette, S. L.; Want, E.; Li, J. V; Barton, R. H.; Boursier-Neyret, C.; Walther, B.; Ebbels, T. M.; et al. Optimized Preprocessing of Ultra-Performance Liquid Chromatography/mass Spectrometry Urinary Metabolic Profiles for Improved Information Recovery. *Analytical Chemistry* **2011**, *83*, 5864–5872.

Vinaixa, M.; Samino, S.; Saez, I.; Duran, J.; Guinovart, J. J.; Yanes, O. A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. *Metabolites* **2012**, *2*, 775–795.

Wall, R.; Ross, R. P.; Fitzgerald, G. F.; Stanton, C. Fatty Acids from Fish: The Anti-Inflammatory Potential of Long-Chain Omega-3 Fatty Acids. *Nutrition Reviews* **2010**, *68*, 280–289.

Walsh, B. H.; Broadhurst, D. I.; Mandal, R.; Wishart, D. S.; Boylan, G. B.; Kenny, L. C.; Murray, D. M. The Metabolomic Profile of Umbilical Cord Blood in Neonatal Hypoxic Ischaemic Encephalopathy. *PLoS ONE* **2012**, *7*, e50520.

Walsh, J. M. E.; Goldberg, J. D. Fetal Aneuploidy Detection by Maternal Plasma DNA Sequencing: A Technology Assessment. *Prenatal Diagnosis* **2013**, *33*, 514–520.

Walsh, J. M.; Wallace, M.; Brennan, L.; McAuliffe, F. M. Early Pregnancy Maternal Urinary Metabolomic Profile and Later Insulin Resistance and Fetal Adiposity. *The Journal of Maternal-Fetal & Neonatal Medicine* **2014**, 1–4.

Wang, Z.; Klipfell, E.; Bennett, B. J.; Koeth, R.; Levison, B. S.; Dugar, B.; Feldstein, A. E.; Britt, E. B.; Fu, X.; Chung, Y.-M.; et al. Gut Flora Metabolism of Phosphatidylcholine Promotes Cardiovascular Disease. *Nature* **2011**, *472*, 57–63.

Want, E. J.; Nordström, A.; Morita, H.; Siuzdak, G. From Exogenous to Endogenous: The Inevitable Imprint of Mass Spectrometry in Metabolomics. *Journal of Proteome Research* **2006a**, *6*, 459–468.

Want, E.; Smith, C.; Qin, C.; Van Horne, K.; Siuzdak, G. Phospholipid Capture Combined with Non-Linear Chromatographic Correction for Improved Serum Metabolite Profiling. *Metabolomics* **2006b**, *2*, 145–154.

Wenk, M. R. The Emerging Field of Lipidomics. *Nature Reviews Drug Discovery* **2005**, *4*, 594–610.

Westerhuis, J. A.; Hoefsloot, H. C. J.; Smit, S.; Vis, D. J.; Smilde, A. K.; van Velzen, E.; van Duijnhoven, J.; van Dorsten, F.; Velzen, E. J. J.; Duijnhoven, J. P. M.; et al. Assessment of PLSDA Cross Validation. *Metabolomics* **2008**, *4*, 81–89.

Whittaker, P. G.; Lind, T. The Intravascular Mass of Albumin during Human Pregnancy: A Serial Study in Normal and Diabetic Women. *British Journal of Obstetrics and Gynaecology* **1993**, *100*, 587–592.

Wiklund, S.; Nilsson, D.; Eriksson, L.; Sjöström, M.; Wold, S.; Faber, K. A Randomization Test for PLS Component Selection. *Journal of Chemometrics* **2007**, *21*, 427–439.

Willker, W.; Leibfritz, D. Assignment of Mono- and Polyunsaturated Fatty Acids in Lipids of Tissues and Body Fluids. *Magnetic Resonance in Chemistry* **1998**, *36*, S79–S84.

Wilson, K.; Hawken, S.; Ducharme, R.; Potter, B. K.; Little, J.; Thébaud, B.; Chakraborty, P. Metabolomics of Prematurity: Analysis of Patterns of Amino Acids, Enzymes, and Endocrine Markers by Categories of Gestational Age. *Pediatric Research* **2014**, *75*, 367–373.

Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; et al. HMDB: The Human Metabolome Database. *Nucleic Acids Research* **2007**, *35*, D521–D526.

Wiznitzer, A.; Mayer, A.; Novack, V.; Sheiner, E.; Gilutz, H.; Malhotra, A.; Novack, L. Association of Lipid Levels during Gestation with Preeclampsia and Gestational Diabetes Mellitus: A Population-Based Study. *American Journal of Obstetrics and Gynecology* **2009**, *201*, 8.

Wu, D. H.; Chen, A. D.; Johnson, C. S. An Improved Diffusion-Ordered Spectroscopy Experiment Incorporating Bipolar-Gradient Pulses. *Journal of Magnetic Resonance, Series A* **1995**, *115*, 260–264.

Xia, J.; Broadhurst, D. I.; Wilson, M.; Wishart, D. S. Translational Biomarker Discovery in Clinical Metabolomics: An Introductory Tutorial. *Metabolomics* **2013**, *9*, 280–299.

Xia, Y.-Q.; Jemal, M. Phospholipids in Liquid Chromatography/mass Spectrometry Bioanalysis: Comparison of Three Tandem Mass Spectrometric Techniques for Monitoring Plasma Phospholipids, the Effect of Mobile Phase Composition on Phospholipids Elution and the Association of Phos. *Rapid Communications in Mass Spectrometry* **2009**, *23*, 2125–2138.

Xu, Q.-S.; Liang, Y.-Z. Monte Carlo Cross Validation. *Chemometrics and Intelligent Laboratory Systems* **2001**, *56*, 1–11.

Yin, P.; Peter, A.; Franken, H.; Zhao, X.; Neukamm, S. S.; Rosenbaum, L.; Lucio, M.; Zell, A.; Häring, H.-U.; Xu, G.; et al. Preanalytical Aspects and Sample Quality Assessment in Metabolomics Studies of Human Blood. *Clinical Chemistry* **2013**, *59*, 833–845.

Young, B. C.; Levine, R. J.; Karumanchi, S. A. Pathogenesis of Preeclampsia. *Annual Review of Pathology* **2010**, *5*, 173–192.

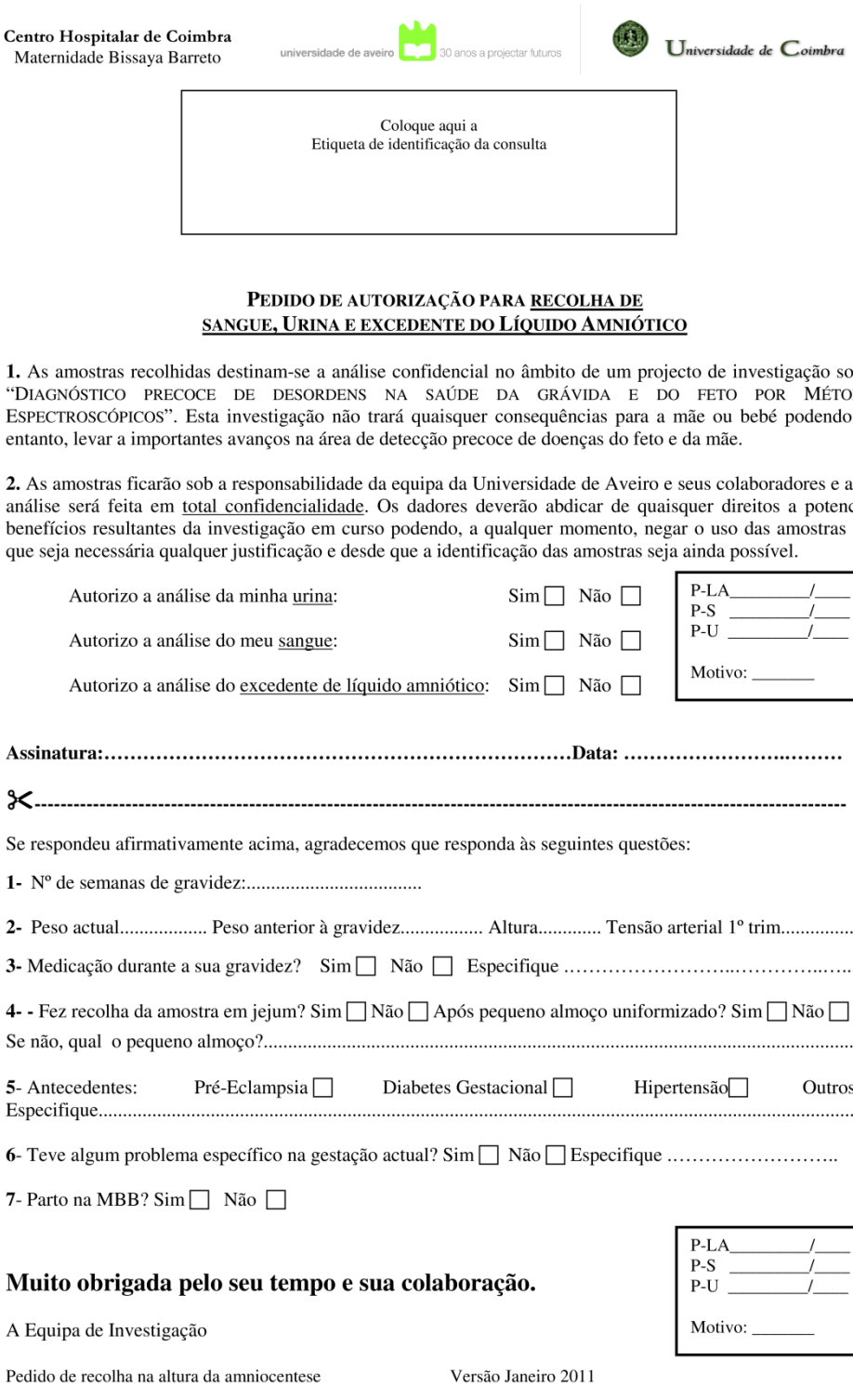
Yu, Z.; Zhai, G.; Singmann, P.; He, Y.; Xu, T.; Prehn, C.; Römisch-Margl, W.; Lattka, E.; Gieger, C.; Soranzo, N.; et al. Human Serum Metabolic Profiles Are Age Dependent. *Aging cell* **2012**, *11*, 960–967.

Yue, B.; Pattison, E.; Roberts, W. L.; Rockwood, A. L.; Danne, O.; Lueders, C.; Möckel, M. Choline in Whole Blood and Plasma: Sample Preparation and Stability. *Clinical Chemistry* **2008**, *54*, 590–593.

Zhao, D.; Han, L.; He, Z.; Zhang, J.; Zhang, Y. Identification of the Plasma Metabolomics as Early Diagnostic Markers between Biliary Atresia and Neonatal Hepatitis Syndrome. *PLoS ONE* **2014**, *9*, e85694.

Zheng, X. X.; Su, M.; Pei, L.; Zhang, T.; Ma, X.; Qiu, Y.; Xia, H.; Wang, F.; Zheng, X. X.; Gu, X.; et al. Metabolic Signature of Pregnant Women with Neural Tube Defects in Offspring. *Journal of Proteome Research* **2011**, *10*, 4845–4854.

Annex: Informed consents and questionnaires used in this study



Centro Hospitalar de Coimbra
Maternidade Bissaya Barreto

universidade de aveiro 30 anos a projectar futuros

Universidade de Coimbra

Coloque aqui a
Etiqueta de identificação da consulta

PEDIDO DE AUTORIZAÇÃO PARA RECOLHA DE SANGUE, URINA E EXCEDENTE DO LÍQUIDO AMNIÓTICO

1. As amostras recolhidas destinam-se a análise confidencial no âmbito de um projecto de investigação sobre: "DIAGNÓSTICO PRECOCE DE DESORDENS NA SAÚDE DA GRÁVIDA E DO FETO POR MÉTODOS ESPECTROSCÓPICOS". Esta investigação não trará quaisquer consequências para a mãe ou bebé podendo, no entanto, levar a importantes avanços na área de detecção precoce de doenças do feto e da mãe.

2. As amostras ficarão sob a responsabilidade da equipa da Universidade de Aveiro e seus colaboradores e a sua análise será feita em total confidencialidade. Os dadores deverão abdicar de quaisquer direitos a potenciais benefícios resultantes da investigação em curso podendo, a qualquer momento, negar o uso das amostras sem que seja necessária qualquer justificação e desde que a identificação das amostras seja ainda possível.

Autorizo a análise da minha urina: Sim ☐ Não ☐

Autorizo a análise do meu sangue: Sim ☐ Não ☐

Autorizo a análise do excedente de líquido amniótico: Sim ☐ Não ☐

P-LA_____/____
P-S _____/____
P-U _____/____

Motivo: _____

Assinatura:……Data: ……..……

✂----------

Se respondeu afirmativamente acima, agradecemos que responda às seguintes questões:

1- Nº de semanas de gravidez:.....

2- Peso actual..... Peso anterior à gravidez..... Altura..... Tensão arterial 1º trim.....

3- Medicação durante a sua gravidez? Sim ☐ Não ☐ Especifique ……..……..…....

4- - Fez recolha da amostra em jejum? Sim ☐ Não ☐ Após pequeno almoço uniformizado? Sim ☐ Não ☐
Se não, qual o pequeno almoço?.....

5- Antecedentes: Pré-Eclampsia ☐ Diabetes Gestacional ☐ Hipertensão ☐ Outros ☐
Especifique.....

6- Teve algum problema específico na gestação actual? Sim ☐ Não ☐ Especifique ……..

7- Parto na MBB? Sim ☐ Não ☐

P-LA_____/____
P-S _____/____
P-U _____/____

Motivo: _____

Muito obrigada pelo seu tempo e sua colaboração.

A Equipa de Investigação

Pedido de recolha na altura da amniocentese Versão Janeiro 2011

Figure A.1. Informed consent given to pregnant women before sample collection at the time of amniocentesis.

Para senhoras ACOMPANHADAS
durante a gravidez

Centro Hospitalar de Coimbra

universidade de aveiro 30 anos a projectar futuros

Versão Jan11

Universidade de Coimbra

Coloque aqui a
Vinheta de identificação da consulta

PEDIDO DE AUTORIZAÇÃO PARA ANÁLISE DE: **- URINA DA MÃE E DO BEBÉ**
- SANGUE DA MÃE

a) As amostras recolhidas destinam-se a análise confidencial para um projecto de investigação sobre: "DIAGNÓSTICO PRECOCE DE DESORDENS NA SAÚDE DA GRÁVIDA, DO FETO E DO RECÉM-NASCIDO POR MÉTODOS ESPECTROSCÓPICOS". Esta investigação não trará quaisquer consequências para mãe ou bebé podendo, no entanto, levar a importantes avanços na detecção precoce de doenças da grávida e do bebé.

b) A urina e, quando possível, o sangue da mãe deverão ser recolhidos nas seguintes alturas:

1. Em cada consulta no Centro Hospitalar, durante a gravidez
2. Na altura da amniocentese, se aplicável,
3. Na altura do parto.

c) A urina do bebé deverá ser recolhida durante a sua estadia no Centro Hospitalar.

d) Informam-se os pais que serão contactados por correio ou email, durante os 2 primeiros anos de vida do bebé, para recolha de informação sobre a saúde do bebé e sua correlação com a análise de urina.

e) As amostras recolhidas ficarão sob a responsabilidade da equipa da Universidade de Aveiro e seus colaboradores e a sua análise será feita em total confidencialidade. Os dadores deverão abdicar de quaisquer direitos a potenciais benefícios resultantes da investigação em curso podendo, a qualquer momento, negar o uso das amostras sem que seja necessária qualquer justificação e desde que a identificação das amostras seja ainda possível.

Autorizo a análise da minha urina: Sim ☐ Não ☐
Autorizo a análise do meu sangue: Sim ☐ Não ☐
Autorizo a análise da urina do meu bebé: Sim ☐ Não ☐
Tomei conhecimento das condições acima descritas, as quais aceito: Sim ☐ Não ☐

Assinatura da mãe:……Data: ……..

✂----------

3º Trimestre

Nº do Processo da mãe..... **Data da recolha:**.....

1- Nº de semanas de gravidez:..... Peso actual..... Altura..... Tensão arterial 3º trim.....

2- Medicação? Sim ☐ Não☐ ..... Problemas durante gravidez? Sim ☐ Não ☐.....

3- Fez recolha da amostra em jejum? Sim ☐ Não ☐ Após pequeno almoço uniformizado? Sim ☐ Não ☐

se não, qual o pequeno almoço?.....

✂----------

2º Trimestre

Nº do Processo da mãe..... **Data da recolha:**.....

1- Nº de semanas de gravidez:..... Peso actual..... Altura..... Tensão arterial 2º trim.....

2- Medicação? Sim ☐ Não☐ ..... Problemas durante gravidez? Sim☐ Não☐ .....

3- Fez recolha da amostra em jejum? Sim ☐ Não ☐ Após pequeno almoço uniformizado? Sim ☐ Não ☐

se não, qual o pequeno almoço?.....

✂----------

1º Trimestre

Nº do Processo da mãe..... **Data da recolha:**.....

1- Nº de semanas de gravidez:..... Peso actual..... Altura..... Tensão arterial 1º trim.....

2- Medicação? Sim ☐ Não☐ ..... Problemas durante gravidez? Sim☐ Não☐ .....

3- Fez recolha da amostra em jejum? Sim ☐ Não ☐ Após pequeno almoço uniformizado? Sim ☐ Não ☐

se não, qual o pequeno almoço?.....

Figure A.1. Informed consent given to pregnant women before sample collection at each pregnancy trimester.